

DEVELOPMENT OF A CRASH RISK PROBABILITY MODEL AND ITS APPLICATION TO MITIGATE HAZARDOUS CONDITIONS ON FREEWAYS

A thesis submitted in fulfilment of the requirements for the degree of
Master of Engineering

Md. Mahmud Hasan

B.Sc.

School of Civil, Environmental and Chemical Engineering

College of Science, Engineering and Health

RMIT University

February 2012

DECLARATION

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work, paid or unpaid, carried out by a third party is acknowledged; and, Ethics procedures and guidelines have been followed.

.....

(Md. Mahmud Hasan)

23.02.2012

ACKNOWLEDGEMENTS

The first and the foremost, the author would like to thank the Almighty for allowing him to complete the research work successfully.

The author expresses his gratitude to a number of people who played an important role in completing the research work. The author is greatly indebted to his supervisor Dr. Shamas Bajwa for his invaluable and expert guidance throughout this research timeline. Special thanks to Dr. Edward Chung for his helpful suggestions. Thanks to Mr. Jude Jusayan and Tim Strickland of Vic Roads and the authority of Bureau of Meteorology (Australia) for providing the data. The author is grateful to Dr. Indubhushan Patnaikuni and Dr. Muhammed Bhuiyan for their kind help during the study period.

The author extends gratitude to all his friends in RMIT University for their help and co-operation. Also, thanks to the office staffs of the School of Civil, Environmental and Chemical Engineering (RMIT University) for providing information and suggestion during this study. The author also appreciates the assistance and supports of Nazmun Nahar.

Finally, the author is grateful to his parents: Humayun Kabir and Nurun Nahar, without whom this study would have been impossible. The author is also thankful to his siblings: Azmin, Razon and Ivan to provide constant support during the hardest time.

TABLE OF CONTENTS

DECLARATION.....	i
ACKNOWLEDGEMENTS.....	ii
LIST OF FIGURES.....	viii
LIST OF TABLES.....	xi
ABSTRACT.....	1
 Chapter 1: INTRODUCTION.....	 2
1.1 Background.....	2
1.2 Statement of the Problem.....	3
1.3 Objectives of the Study.....	4
1.4 Research Questions.....	4
1.5 Outline of the Thesis.....	4
 Chapter 2: LITERATURE REVIEW.....	 6
2.1 Effect of Traffic Parameters on Road Accident.....	6
2.1.1 Traffic Flow and Road Accidents.....	6
2.1.2 Vehicle Speed and Road Accidents.....	7
2.2 Effect of Rainfall on Traffic Parameters and Road Accidents.....	8
2.2.1 Rainfall and Traffic Flow.....	8
2.2.2 Rainfall and Vehicle Speed.....	9

2.2.3 Rainfall and Road Accidents.....	10
2.3 Accident Prediction Models.....	11
2.4 Application of Mitigation Schemes.....	12
2.5 Limitations and Research Gaps.....	13
2.6 Summary.....	14
 Chapter 3: SITE SELECTION AND DATA ISSUES.....	 15
3.1 Selected Site.....	15
3.2 Data Requirements.....	16
3.3 Data Collection.....	16
3.3.1 Data Collection Site.....	16
3.3.2 Duration of Data.....	17
3.3.3 Details of Each Data Type.....	17
3.4 Data Fusion.....	19
3.5 Summary.....	20
 Chapter 4: MODEL DEVELOPMENT AND VALIDATION.....	 21
4.1 Classification Tree Analysis.....	21
4.1.1 Building the Classification Tree.....	21
4.1.2 Splitting Criteria.....	22
4.1.3 Missing Data Cases.....	23
4.1.3.1 Choosing the Split.....	23
4.1.3.2 Surrogate Variables.....	24

4.2 Formulation of Classification Tree Model for this Study.....	24
4.2.1 Selection of Model Variables.....	24
4.2.2 Building of Classification Tree Model.....	26
4.2.3 Cross-Validation Result.....	28
4.2.4 Pruned Classification Tree.....	29
4.2.5 Crash Hazard Prediction Model Based on Classification Tree.....	30
4.2.6 Model Validation.....	31
4.3 Hazard Prediction Index.....	34
4.3.1 Variables Used for Developing the Hazard Prediction Index.....	34
4.3.2 Model Formulation.....	35
4.3.3 Selected Hazard Prediction Index.....	39
4.3.4 Validation of the Indices.....	41
4.3.5 Selection of the Best Accident Prediction Index.....	42
4.4 Comparative Analysis of Classification Tree and Accident Prediction Model.....	42
4.5 Selected Crash Risk Probability Model for the Study.....	44
4.6 Summary.....	46
 Chapter 5: SIMULATION STUDY.....	 47
5.1 Simulation Features.....	47
5.2 Input Data for Simulation.....	48
5.2.1 Network Layout.....	48
5.2.2 Traffic Demand Data.....	48
5.2.2.1 OD Matrix for Outbound Direction.....	48

5.2.2.2 OD Matrix for Inbound Direction.....	49
5.2.3 Vehicle Configuration.....	50
5.3 Calibration Procedure for Simulation Output.....	50
5.3.1 Selected Parameter for Calibration.....	51
5.3.2 Origin-Destination Flow Estimation.....	52
5.3.3 Calculating the Speed of a Vehicle in a Section.....	53
5.3.4 Calibration Results.....	55
5.4 Validation Process for Simulation Output.....	57
5.4.1 Selected Parameter for Validation.....	58
5.4.2 Validation Results.....	59
5.5 Summary.....	61
 Chapter 6: APPLICATION OF MITIGATION SCHEMES FOR HAZARDOUS CONDITION.....	 62
6.1 Reasons for Applying Mitigation Scheme.....	62
6.2 Different Mitigation Schemes.....	63
6.3 Selected mitigation schemes.....	64
6.3.1 Ramp Metering.....	64
6.3.2 Variable Speed Limit (VSL).....	65
6.3.3 Effectiveness of the Selected Mitigation Strategies.....	66
6.4 Application of Mitigation Scheme in Simulation Software (AIMSUN).....	66
6.4.1 Application of Ramp Metering in AIMSUN.....	66
6.4.1.1 Types of Ramp Metering.....	66

6.4.1.2 Types of the Control.....	67
6.4.1.3 Parameter Selection.....	67
6.4.2 Application of Variable Speed Limit System in AIMSUN.....	68
6.5 Traffic Scenarios Analysis Method.....	68
6.6 Application of Ramp Metering Scheme.....	70
6.7 Application of Variable Speed Limit (VSL) Strategy.....	71
6.8 Parameter used for ramp metering and VSL strategy.....	72
6.9 Effects of mitigation schemes to reduce hazardous situations.....	73
6.10 Selection of the best mitigation scheme.....	90
6.11 Summary.....	94
 Chapter 7: CONCLUSIONS.....	 95
7.1 Research Summary and Contributions.....	95
7.2 Future Research Directions.....	96
 REFERENCES.....	 98
 APPENDIX A.....	 108
 APPENDIX B.....	 110

LIST OF FIGURES

Figure 3.1: Eastern Freeway in Melbourne (Australia).....	15
Figure 3.2: Required data for this study.....	16
Figure 3.3: Selected weather stations and the test bed.....	17
Figure 3.4: Flow-chart of data cleansing process.....	18
Figure 3.5: Flow-chart of data fusion.....	20
Figure 4.1: Total Classification tree.....	27
Figure 4.2: Cross-validation error vs. size of the tree.....	29
Figure 4.3: Pruned classification tree after cross-validation.....	30
Figure 4.4: Comparative graph of predicted and actual value of the accident and non-accident cases.....	32
Figure 4.5: Accident prediction index formulation steps.....	36
Figure 5.1: Origins and destinations for outbound sections of Eastern freeway.....	48
Figure 5.2: Origins and destinations for inbound sections of Eastern freeway.....	49
Figure 5.3: Framework for calibration process.....	51
Figure 5.4: OD matrix estimation process.....	52
Figure 5.5: Parameters used for selected vehicle types.....	55
Figure 5.6: Calibration result for outbound sections on 29/10/2009.....	56
Figure 5.7: Calibration result for inbound sections on 23/05/2009.....	57
Figure 5.8: Framework for validation process.....	58
Figure 5.9: Validation result for outbound sections on 16/11/2009.....	60
Figure 5.10: Validation result for inbound sections on 11/06/2009.....	61
Figure 6.1: Scenario analysis steps.....	68

Figure 6.2: Selection of parameters for Ramp Metering strategy.....	70
Figure 6.3: Selection of parameters for Variable Speed Limit strategy.....	71
Figure 6.4: Time-space diagram of hazardous condition during no-control situation in peak hour traffic.....	74
Figure 6.5: Improvements of hazards by ramp metering schemes in peak hour traffic.....	76
Figure 6.6: Time-space diagram of hazardous condition by ramp metering in peak hour traffic.....	76
Figure 6.7: Improvements of hazards after VSL schemes in peak hour traffic.....	78
Figure 6.8: Time-space diagram of hazardous condition by VSL in peak hour traffic.....	78
Figure 6.9: Time-space diagram of hazardous condition during no-control situation in low volume traffic.....	80
Figure 6.10: Improvements of hazards by ramp metering schemes in low volume traffic.....	81
Figure 6.11: Time-space diagram of hazardous condition by ramp metering in low volume traffic.....	82
Figure 6.12: Improvements of hazards after VSL schemes in low volume traffic.....	83
Figure 6.13: Time-space diagram of hazardous condition by VSL in low volume traffic.....	84
Figure 6.14: Time-space diagram of hazardous condition during no-control situation in high volume traffic.....	86

Figure 6.15: Improvements of hazards by ramp metering schemes in high volume traffic.....	87
Figure 6.16: Time-space diagram of hazardous condition by ramp metering in high volume traffic.....	88
Figure 6.17: Improvements of hazards after VSL schemes in high volume traffic.....	89
Figure 6.18: Time-space diagram of hazardous condition by VSL in high volume traffic.....	90
Figure 6.19: Comparison of the improvements by two schemes in peak hour traffic.....	91
Figure 6.20: Comparison of the improvements by two schemes in low volume traffic.....	91
Figure 6.21: Comparison of the improvements by two schemes in high volume traffic.....	92
Figure B.1: Time-space diagram of hazardous condition of upstream and downstream section where VSL is applied in peak hour traffic.....	110
Figure B.2: Time-space diagram of hazardous condition of upstream and downstream section where VSL is applied in low volume traffic.....	111
Figure B.3: Time-space diagram of hazardous condition of upstream and downstream section where VSL is applied in high volume traffic.....	112

LIST OF TABLES

Table 4.1: CP table (Errors in each split).....	28
Table 4.2: Accident prediction model (chart view).....	31
Table 4.3: Summary of the cases used for modelling and validation process.....	32
Table 4.4: Prediction capability of classification tree.....	33
Table 4.5: Criteria of Hazard prediction indices.....	39
Table 4.6: Prediction capability of the accident prediction indices.....	42
Table 4.7: Comparison of prediction capability of two selected models.....	43
Table 5.1: Framework for OD matrix in outbound direction.....	49
Table 5.2: Framework for OD matrix in inbound direction.....	50
Table 5.3: Selected OD table.....	53
Table 6.1: Improvements of hazard condition after ramp metering in peak hour traffic.....	75
Table 6.2: Improvements of hazard condition after VSL schemes in peak hour traffic.....	77
Table 6.3: Improvements in hazards by ramp metering schemes in low volume traffic.....	80
Table 6.4: Improvements in hazard condition after VSL schemes in low volume traffic.....	82
Table 6.5: Improvements of hazard condition by ramp metering in high volume traffic.....	86
Table 6.6: Improvements of hazard condition after VSL schemes in high volume traffic.....	88
Table 6.7: Selected mitigation schemes in three different traffic conditions.....	93

Table 6.8: Conditions of hazard at upstream and downstream section before and after applying selected mitigation schemes.....	94
Table A.1: Features of traffic data.....	108
Table A.2: Condition of accident data.....	109
Table A.3: Condition of weather data.....	109

ABSTRACT

This research presents a method for the identification of hazardous situations on the freeways as well as the hazard mitigation system based on the identification. The hazard identification is done using a crash risk probability model and based on the identification mitigation strategy is employed to reduce the effect of the hazard. For this study, about 18 km long section of Eastern Freeway in Melbourne, Australia is selected as a test bed. Three categories of data i.e. traffic weather and accident record data are used for the analysis and modelling.

In developing the crash risk probability model, initially, two models i.e. classification tree and Hazard Prediction Index are developed in this study. In formulating the models, it is found that weather conditions do not have significant impact on accident occurrence so both the models are formulated using two traffic indices; traffic flow and vehicle speed only. The classification tree consists of twenty eight conditions which identify the hazard and non-hazard situations. On the other hand, Hazard Prediction Index is formulated by the differences of traffic parameters with threshold values. Seven different prediction indices are examined and best one is selected as Hazard prediction Index model based on prediction error minimisation. After comparing the prediction capabilities of the two models i.e. classification tree and Hazard Prediction Index, latter is chosen as the crash risk probability model for this study.

Two mitigation schemes i.e. Ramp Metering and Variable Speed Limit have been selected in this study for conducting scenario analysis to reduce the hazard risks on freeways. These mitigation schemes are tested for three different traffic conditions i.e. low traffic volume, peak hour traffic and high traffic volume in order to find out the best mitigation scheme. Percentage reduction in Hazard Prediction Index is used as the selection criterion for the best mitigation scheme under different traffic conditions. After the evaluation, it is observed that Variable Speed Limit scheme is more effective than ramp metering scheme in order to mitigate hazardous situations under different traffic scenarios on the freeway.

Chapter 1

INTRODUCTION

1.1 Background

Based on user convenience, five most popular modes of transport operated throughout the world are roadway, railway, inland water, maritime and air transport. Among these modes, the roadway traffic, both for passenger and freight movements, is increasing rapidly. Rapid changes in land use pattern in conjunction with socio-economic development accelerate the demand for effective roadway network and services. Therefore, a sound road traffic service is vital for any road network for the safe, efficient and faster movements of people and goods to their destinations.

Traffic safety can be impaired for several reasons which can be classified into three broad categories which are

1. Driver factors e.g. driver fatigue, attention, drink driving
2. Vehicle factors e.g. mechanical faults resulting in out of control vehicles
3. Environmental factors e.g. traffic conditions on the road as well as weather

Driving factors such as fatigue, excessive drugs or alcohol consumption, training of drivers, pressure to save time etc. cause road incidents. All these matters distract driver's attention and reduce driving efficiency (Aljanahi et al., 2009). Vehicle factors consisting of manufacturing or mechanical faults also result road incidents. These factors include brake failure, tyre blow out, steering mechanism failure etc. Similarly, weather conditions such as rainfall, snowfall, fog or smog etc. have effects on road safety. These weather events induce crash risk by reducing driver visibility, decreasing road pavement friction or by lane obstruction, submersion (Goodwin, 2002).

Among these three categories, only environmental factor is considered in this study. As an environmental factor, traffic conditions on the road can cause accidents. Traffic performance indicators such as traffic flow and speed which can act as a proxy for the traffic condition may

indicate hazardous situations leading to accidents. After the accident, these parameters are altered dramatically as accidents usually disrupt the normal flow of traffic resulting in congestion. Weather conditions may also lead to situations which might hamper usual traffic movements and roadway safety. Adverse weather conditions like rainfall and snowfall not only reduce drivers' visibility but also make the roads dangerous because of reduced friction between tyres and road surfaces. For improving road safety and traffic performance, many countries have adopted different initiatives targeting different categories of accident causing factors. These include driver education programs and graduated licensing systems to reduce driver factors, vehicle fitness certifications to reduce the accidents caused by vehicle faults and traffic management plans e.g. tidal flow operation, exclusive lanes etc. and traffic control strategies e.g. ramp metering, variable speed limit, co-ordinated traffic signal etc.

1.2 Statement of the Problem

Road accidents cause vehicle damage and serious human injury even deaths. According to World Health Organization (WHO), each year approximately 1.2 million people die in road incidents and about 50 million people are injured. Peden (2002) reported that a quarter of all injury deaths are due to road traffic injuries. Road accident also induces traffic congestion. Due to accident, some or all the lanes of the roadway are blocked and vehicles are stuck in congestion for long durations. This congestion results in increased travel time, vehicle emissions and fuel usage etc. (Golob et al., 2004b).

Some of the previous studies have investigated the impacts of traffic parameters (vehicle speed and traffic flow) and weather on road accidents, but fewer studies have developed mitigation procedures to reduce the crash risk probability. This study not only develops the relationship among traffic parameters, weather and crash risk but also explores possible traffic control strategies in order to mitigate the hazardous condition.

The outcomes of this research are general, and can be applied to similar roads in any area as the selected site and data are in accordance with global standards.

1.3 Objectives of the Study

The scope of this research is limited to investigating the impact of environmental factors on crash risks and to develop mitigation schemes to reduce these risks while maintaining the traffic performance indicators.

The main objectives of the study are:

- To establish the relationship among traffic performance indicators, weather condition and road accidents.
- To develop a model to predict the crash risk probabilities depending on traffic Condition.
- To develop an effective traffic control strategy to reduce the hazardous condition caused by traffic condition on the roadway.

1.4 Research Questions

The research questions for this study are:

- Does the roadway crash risk depend on traffic parameters (flow and speed) and weather condition?
- How can the road hazards be correlated with traffic flow, vehicle speed and weather condition?
- What will be the plausible roadway crash risk mitigation schemes?
- What are some effective traffic control strategies to lessen the crash risks on a roadway?

1.5 Outline of the Thesis

Apart from this chapter, the rest of this thesis consists of seven chapters. The chapters are organised as follows:

Chapter 2 reviews the existing literature on the proposed research topic. A critical review of the existing research identifies the gaps in the existing body of knowledge.

Chapter 3 discusses issues such as data requirements, collection of data and arrangement of the dataset. This chapter also describes the selected site and various features of the selected roadway.

Chapter 4 presents the crash risk model and also lays out the procedure for variable selection, missing data issues, cross-validation, and shows the result of the validation of the model.

Chapter 5 outlines the simulation process used in this study to verify the effectiveness of the proposed crash risk model. This chapter focuses on the calibration and validation of the created simulation model for this study.

Chapter 6 discusses the accident mitigation schemes. This chapter emphasises two types of traffic control strategies: ramp metering and variable speed limit; and shows their effectiveness to reduce the hazardous condition on the roadway. Based on the comparison of the effectiveness this chapter provides a methodology to select the best mitigation scheme for different scenarios.

Chapter 7 gives the summary and conclusions of the study as well as some recommendations for improvements in the suggested scheme. This chapter also suggests directions for future research.

Chapter 2

LITERATURE REVIEW

Literature review is one of the elementary parts of this study. The objectives and research questions of this study are selected after the critical review of the relevant existing literature. This chapter reviews and summarises the previous studies investigating the impact of environmental factors including traffic conditions and weather on the road accidents. Existing crash risk and accident prediction models are also reviewed. A review of the existing traffic mitigation schemes which can be used to alleviate the crash risk is also provided. The chapter closes with a critique of the gaps in the existing body of knowledge which highlight the importance of this research

2.1 Effect of Traffic Parameters on Road Accident

2.1.1 Traffic Flow and Road Accidents

A number of studies have been conducted to relate traffic flow with road crash propensity. Martin (2002) showed that damage-only and injury-involved incident rates are higher in light traffic than in heavy traffic conditions. He also compared the incident rates on the basis of time of the day and found that these rates do not depend on day time or night-time traffic. Golob and Recker (2004) also found out the correlation between traffic flow and crash risk. They developed twenty one traffic flow regimes which have different profiles for different road crash types. Their study is limited to the urban freeways consisting of three or more lanes in each direction. The findings of this study cannot be confirmed to have the accuracy for spatial transferability because of missing data and lack of perfection in calculation of flow rates. Hasan et al. (2011) have found that road accident probability on the freeway depends significantly on the traffic flow. They developed a regression tree by using traffic flow and speed at accident location, nearest upstream and downstream and concluded that road accidents depend more on traffic parameters: traffic flow and vehicle speed rather than weather condition.

On the contrary, Lord et al. (2005) showed that the crash risk cannot be predicted perfectly only by traffic flow but adding traffic density so improves the prediction performance. Furthermore, they also described the comparative difference of crash-density relationship between urban and rural freeways. For the same flow and density, it has been found that crash rates are much higher on urban freeways than the rural ones.

Dickerson et al. (1998) revealed significant differences in accident- traffic flow relationship by road class and geography. Their outcomes are based on all types of accidents regardless of severity level. Accident probability also depends on type of vehicles (Ayati and Abbasi, 2011). Non-passenger car vehicles are found to cause more accidents on urban highways than other vehicle types. Interestingly, it was shown that heavy vehicles cause less accident than non-passenger cars including taxis and motorcycles.

In order to relate traffic parameters and road geometry with different types of accidents, Christoforou et al. (2011) found that two-vehicle side swipe accident probability increases with traffic volume, however, multiple vehicle side swipe crash is more dependent on higher speed. They also mentioned that road geometry and speed play a significant role to cause single vehicle accident and multi-vehicle crash respectively. They did not find any significant effect of weather and standard deviation of traffic on road crashes. Average Annual Daily Traffic (AADT) volume is also found to be an important factor directly related to number of road accidents (Karlaftis et al., 2002; Wang et al. 2009).

2.1.2 Vehicle Speed and Road Accidents

Aarts et al. (2006) reviewed the literature on vehicle speed and road accident relationship and showed that road incidents increase significantly with an increase in speed on minor roads than on major roads. Similarly, Navon (2001) mentioned that excessive speed causes road crashes. He also described that the relationship between average speed and accident is not clear.

Taylor et al. (2002) developed a speed-accident relationship for single carriageway roads in England. They showed that all categories of accident increased rapidly with the mean speed and indicated that a 10 % increase in mean speed results in a 26% increase in the frequency of all injury accidents. Relationship between fatal accident and speed was found to be

significant. Similarly, Elvik (2004) described that mean speed is positively related with frequency and severity of accidents since the number of road accidents increases with increase in speed. In order to develop the relationship between speed and safety, Aljanahi et al. (1999) found significant positive relationship between mean speed and accident rate. Similarly, Hasan et al. (2011) also concluded that vehicle speed is a contributing factor to accidents' occurrence on freeway.

Waaed et al. (1994) suggested that four lane motorways are relatively safer than other road types. Also, Taylor et al. (2000) discussed the results of driver based and road based previous studies, and mentioned that higher speed causes more accidents. The outcome of the study showed that approximately 1 km/hr decrease in average speed can reduce accident occurrence rate by 2% - 7%.

In order to relate the speed limit and fatal crash, Ossaiaander (2002) found that in Washington (USA) speed limit and fatal crash occurrence have positive relationship, higher the speed more the number of fatal accidents. But the relationship between speed limit and all types of accident rate is not so clear.

While aiming to assess the effects of traffic congestion on the frequency of crash rate by using spatial analysis approach, Wang et al. (2009) found that traffic congestion has no impact on the frequency of accident occurrence (either for fatal crash or slight injury crash).

2.2 Effect of Rainfall on Traffic Parameters and Road Accidents

2.2.1 Rainfall and Traffic Flow

Most of the existing literature shows that rainfall, in general, reduces traffic flow. Cools et al. (2008) found that rainfall has a clear decreasing effect on traffic flow both in upstream and downstream locations in Belgian coastline areas although the exact reduction percentage is not mentioned. Chung et al. (2005) revealed that on Tokyo Metropolitan Expressway (MEX), the capacity decreases are in the range of 4% to 14% with increasing rainfall intensity; and this reduction percentage was quite similar to the findings of Maze et al. (2006) and Hranac et al. (2007) which were based on the freeways in United States. Hogema (1996) did not get any

significant reduction in daily traffic volume due to rain on motorway near the city of Breda (Netherlands), however, the distance between weather station and roadway was large.

Agarwal et al. (2005) showed that rainfall impact on urban freeway traffic operations were different from the recommended reduction percentage for traffic flow mentioned in the Highway Capacity Manual (HCM 2000). This research found that light rain had a significant effect (5% - 10% reduction) on capacity, whereas, this impact was not mentioned in HCM 2000. HCM 2000 mentions that heavy rain decreases capacity by 14% – 15% while Smith et al. (2004) and Prevedouros and Kongsil (2003) found that heavy rainfall decreased capacity by 25% - 30% and about 20% respectively. Furthermore, they also got significant reduction percentages in capacity due to light rain which was not considered in HCM 2000. Keay and Simmonds (2005) investigated the impact of rainfall on traffic volume at two freeways in Melbourne (Australia) both in day and night periods, and found a negative relationship between volume and rainfall amount i.e. higher the rain intensities, larger the reduction in traffic volume.

The effect of rainfall is different on different days of the week. Chung et al. (2005), Al Hassan and Barker (1999) and Changnon (1996) revealed that compared to weekdays, weekends traffic is more sensitive to weather conditions although their study areas were different.

2.2.2 Rainfall and Vehicle Speed

Saberi and Bertini (2010) showed that the average speeds during no rain period were up to 6.5 km/hr higher than during rainy conditions in Oregon (USA). Billot (2010) showed that on a slow lane, there was a clear decline of the frequencies of free flow speeds in rainy conditions while the frequency of speeds in the range of 70 - 90 km/hr were higher under light and medium rain than no rain condition. Similarly, Hranac et al. (2007) revealed that light rain decreased free flow speed and speed at capacity by about 2% - 4% and 8% - 10% respectively, and reduction increases with the rain intensity.

Kyte et al. (2000) found a significantly higher free flow speed reduction due to rainfall than others. They concluded that the effects of light rain and heavy rain on free flow speed may be 50% higher than that presented in HCM 2000. Chung et al. (2006) also found reduction in free flow speed with increasing rainfall intensity and reduction percentages found in this study

nearly matched those found in Smith et al. (2003) albeit their study locations were different. Maze et al (2006). Agarwal et al. (2005) got quite similar percentage (about 1% - 7%) of speed reduction on the freeways. They concluded that reduction in operating speeds on urban freeways during light rainfall was comparable to the HCM 2000 values but for heavy rain, HCM 2000 overstated the speed reduction. This finding contradicts the results presented in Prevedouros and Kongsil (2003).

Rainfall impact on vehicle speed also depends on number of lanes as Brilon and Ponzlet (1996) showed that vehicle speed reduces approximately 10km/hr and 12 km/hr for two and three lane wet freeways. The reduction percentage of vehicle speed in rainy condition was quite higher for arterial roads as Goodwin (2002) found the reduction of speeds on arterial roads was in the range of 10% - 25%. On the contrary, Sabir et al. (2010) showed that rain has no significant effect on commuting car speed.

2.2.3 Rainfall and Road Accidents

Quite a few studies have been conducted to investigate the relationship between road accidents and rainfall. Koetse and Rietveld (2007) did a literature review and concluded that adverse weather (i.e. rainfall) increases the number of road accidents. Similarly, Abdel-Aty and Rajashekar (2006) observed that about 11% of the total number of crashes occurred during rainy periods during 1999 - 2001 in Florida (USA). While doing the accident analysis in the Tokyo Metropolitan Expressway, Chung et al. (2005) found that in the six year period (1998-2004), 5,700 accidents were recorded on rainy days. These results displayed a significant percentage of crash occurrences during rainy conditions and highlighted the need to further investigate the effect of rainfall on road accidents. However, in order to predict accident probability, Hasan et al. (2011) did not find any significant relationship between road accidents and rainfall. Only one year dataset was available for this study and missing data further reduced the explanatory capability of the model.

To relate accidents with traffic flow and rainfall, Agbolosu-Amison et al. (2004) conducted an investigation and found that during rainy conditions, accident rates were higher for traffic flows ranging from 1200 to 1500 veh/hr/ln on freeways and 100 to 200 veh/hr/ln on rural two-lane highways. Keay and Simmonds (2006) found that effect of rain on accident occurrences

also depends on time of the day as considerable differences were found in accident percentages during day and night times on two freeways in Melbourne, Australia. Saberi and Bertini (2010) found that incident rate under moderate rain is higher than light rain conditions.

In order to investigate the rainfall impact on accidents in seven cities of St. Louise metropolitan area, USA, Sherretz and Farhar (1978) observed a positive correlation between accidents and rainy condition but in the analysis of accident severity, such relationship was not found since different results were obtained for different cities. Some of the cities showed negative correlation while others showed positive. Unlike most other studies, Changnon (1996) compared rainfall impact not only with non-rain period but also spatially (city, suburban, rural) in Chicago (USA). His findings show that rainfall influences road crash occurrences. In addition, during rainy conditions, city traffic suffers most compared to suburban and rural traffic. Accident rate comparison may be biased as there are differences in traffic conditions among rural, suburban and city areas.

2.3 Accident Prediction Models

Mustakim et al. (2008) proposed an accident prediction model based on the dataset of Federal Route 50 in Malaysia. In this study, they considered number of access point per kilometre of the roadway, hourly traffic volume, time gap between vehicle and 85th percentile speed. Multiple linear regression model resulted in good accuracy level. Similarly, Hong et al. (2005) also used multiple regression methods in order to develop a crash prediction model but they focused more on road geometry compared to traffic conditions in choosing the independent variables. Road geometry was also considered as predictors in accident prediction model by Kalokota and Seneviratne (1994) but the selected geometry variables are different. Hong et al. (2005) chose number of intersections, connecting roads, pedestrian traffic signals, existence of median barrier, lanes whereas Kalokota and Seneviratne (1994) selected degree of curvature, section length, vertical grade, number of lane, right shoulder width, traffic volume as predictors. This may be due to different site locations i.e. urban and rural highways.

Eisenberg (2004) has developed a crash risk prediction model based on weather variables such as precipitation and snowfall. Negative binomial regression method was used in this

study and accident frequencies were predicted in terms of fatality, injury and property damage only. Similarly, Shankar et al. (1995) made accident frequency prediction model using negative binomial regression. They considered both the road geometry and weather factors to develop the model.

To develop a road incident prediction model for the multilane roads in Italy, Caliendo et al. (2007) took geometric features of roadway and traffic parameters. They revealed that length, curvature and AADT are the significant variables for curved road sections whereas length, number of junctions and AADT are significant for tangent sections of the road. They also showed that negative multinomial regression is better than Poisson regression and negative binomial regression.

Hwan et al. (2005) proposed an accident prediction model for signalised intersections in Seoul City in order to control the local and random types of accidents. They used generalised log-linear model to predict the accident as they have categorical variables in dataset. Similarly, Greibe (2003) developed prediction model based on traffic and geometric variables for urban area. Poisson regression model was used in this study. Pham et al. (2010) developed a model using random forest method by disaggregated traffic data in order to identify the rear end crash on motorway. This study is able to differentiate non-crash and precrash situations by using this methodology. They concluded that speed within a lane need to be regulated in order to reduce rear end crash.

It has been observed from the above discussion that most of existing models are based on aggregated data. Thus the prediction is not for the real time operational model but more for strategic planning model.

2.4 Application of Mitigation Schemes

Bevrani et al. (2011) showed that variable speed limit can reduce the secondary and rear-end crashes. They also described that the reduction of speed on the upstream sections of the congested area increases safety. Lee et al. (2006) found that real time variable speed limit can decrease crash potential by 5% - 17%. Hegyi et al. (2005) discussed that dynamic speed limit can reduce congestion significantly. Nissan and Koutsopoulos (2011) showed that mandatory variable speed limits have more impact on traffic performance and safety than recommended

variable speed limits. Abdel-Aty et al. (2006) concluded that variable speed limits can improve safety and significantly decrease the hazardous condition at a certain location. Similarly, Jiang et al. (2011) revealed that variable speed limit strategy can reduce secondary road crashes during congestion and improve traffic safety.

Luk et al. (2008) mentioned that ramp metering system successfully reduces the duration of the congestion at the bottlenecks by 30% with a corresponding 19% reduction in ramp flow. Taale and Middelham (2000) showed that ramp metering system in the motorways of Netherlands can reduce travel times from 3% - 10%. Lee et al. (2005) found that ramp metering decreases crash potential by 5% - 37% compared to no-control system. The findings of the study done by Abdel-Aty et al. (2007) revealed that the more the ramp can be metered, the more safety benefits can be found. Salem and Papageorgiou (1995) showed that an efficient ramp metering system improves traffic conditions on motorway as well as on arterial roads.

2.5 Limitations and Research Gaps

From the review of the existing literature, it is observed that most of the studies were conducted to (a) relate rainfall with traffic parameters and accidents, (b) find relationship between road geometry and traffic parameters and road crashes, (c) create accident prediction model based on traffic and road geometry.

Extensive research has been carried out to investigate the relationship among rainfall, traffic parameters and road incidents in different regions throughout the world; and most of the studies have similar results.

It is evident from the above review that most of the studies focused either on the development of the accident prediction models based on the aggregated traffic flow and road geometry parameters which can be beneficial for strategic planning purposes but may not be sufficient to mitigate the hazardous traffic situations in real-time. In order to identify the real-time hazardous conditions on the roads, it is necessary that crash hazard prediction models utilising real-time traffic parameters are developed. These models can be used to evaluate different mitigation schemes to reduce the crash hazards in real-time. Not many of the studies have tackled the development of a real-time crash hazard prediction model and its application to

mitigate the hazardous situation so, development of a crash hazard model and its application to mitigate hazardous condition will be an effective contribution to traffic engineering and intelligent transport systems' body of knowledge.

2.6 Summary

This chapter reviews the existing literature investigating effects of traffic parameters and weather on accidents which leads to a better understanding of the need for accident prediction models and their role in developing the mitigation schemes. Gaps in the existing research are identified which helps to formulate the research questions for this study.

Chapter 3

SITE SELECTION AND DATA ISSUES

In order to investigate the impact of traffic conditions on the hazard occurrence and to develop a crash hazard prediction model using real-time traffic performance and accident data, we selected Eastern Freeway in Melbourne as the test bed. This chapter provides details about the selected site, data collection, types and quality of data, data processing and data fusion. Results of preliminary analysis detailing the impact of weather on traffic performance are also presented.

3.1 Selected Site

The selected site for this study, Eastern Freeway in Melbourne (Australia) is one of the important urban freeways for commuting to city from eastern suburbs of Melbourne. The section for the study is approximately 18 km long, from Hoddle Street to Springvale road, consisting of three to five lanes in each direction. Figure 3.1 depicts the selected freeway section.

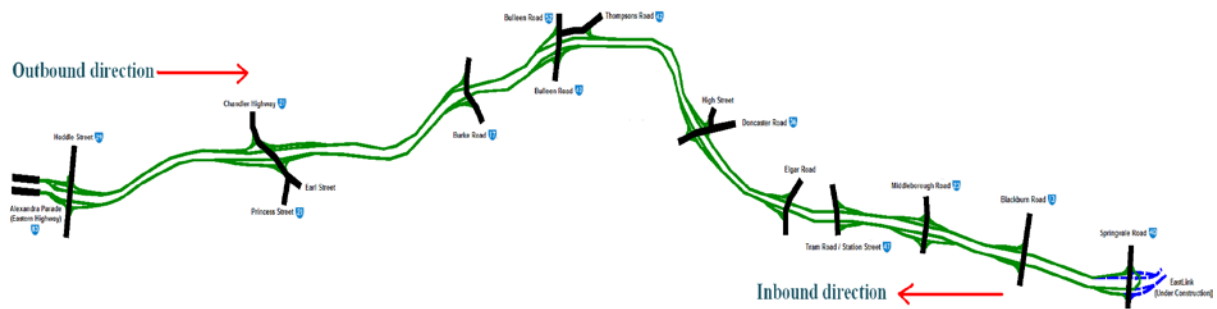


Figure 3.1: Eastern Freeway in Melbourne (Australia) (adapted from Wikipedia (09/02/2012))

Between entry and exit points of the freeway, there are 7 off- ramps and 8 on- ramps in inbound direction and there are 8 off- ramps and 8 on- ramps in the outbound direction. The total roadway section is equipped with 65 detectors in which 32 detectors are located in outbound direction while the remaining 33 detectors are located in inbound direction.

3.2 Data Requirements

Three types of data are required in this study. These are: traffic data, accident data and weather data. Traffic flow and vehicle speed are considered as traffic performance indicator in this study. In case of accident data, detailed information for each of the road incidents including location, time, and severity is required for this study. Rainfall is considered as proxy of the adverse weather conditions hence rainfall intensity data are also needed in this study to investigate the impact of rainfall on accidents. Figure 3.2 depicts an overview of data requirements.

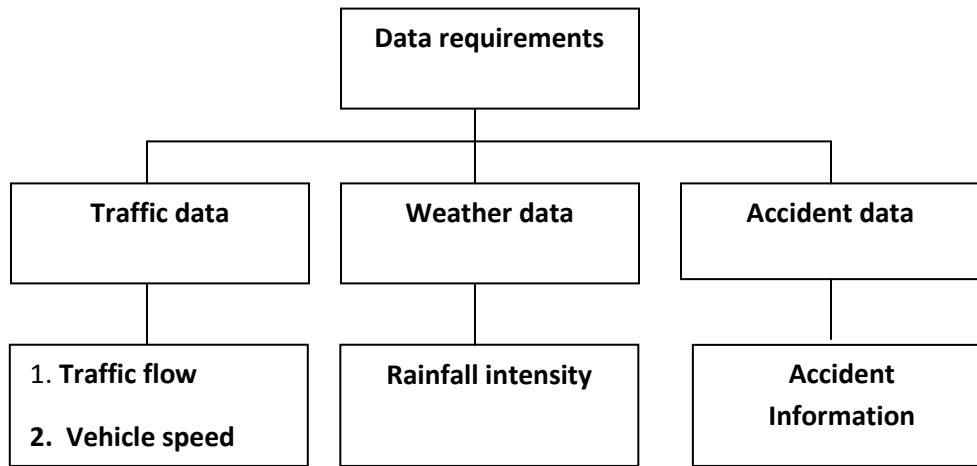


Figure 3.2: Required data for this study

3.3 Data Collection

3.3.1 Data Collection Site

The required data are collected from two different organisations provided Traffic data used in this study were collected by VicRoads which is the road authority of Victoria (Australia) from the loop detectors on the selected roadway. VicRoads also provided detailed road accident information. Although the traffic data and crash data are provided by the same authority, these two datasets are distinct and need to be matched together into a new database in order to be used for analysis. Bureau of Meteorology (Australia) is responsible for collecting rainfall data from three weather stations near the selected roadway. These three weather stations are View

bank (Apsana), Scoresby Research Institute and Melbourne regional office. Figure 3.3 shows the locations of the three weather stations as well as the selected freeway.

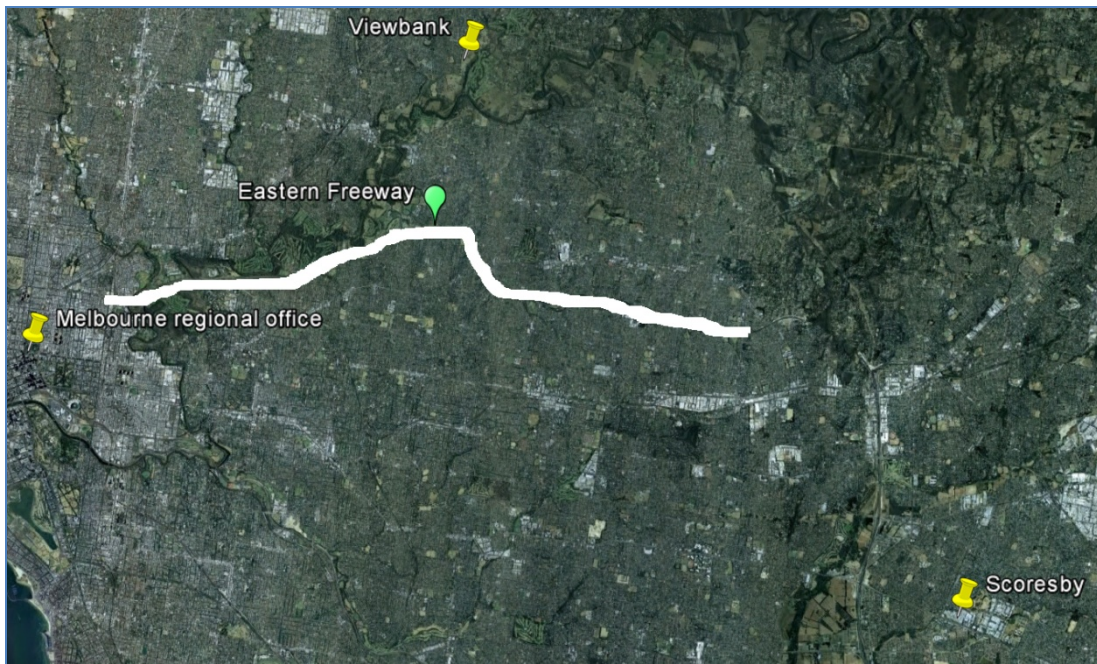


Figure 3.3: Selected weather stations and the test bed

3.3.2 Duration of Data

Traffic, rainfall and accidents data used in this study are from September 2007 till June 2010 which is approximately three years. Actual data for different sources are for longer durations but continuously overlapping data for all three categories was only available from September 2007 till June 2010

3.3.3 Details of Each Data Type

Traffic data: Traffic detector data are collected and stored every minute as raw data i.e. in binary format which is stored in different files based on temporal sequence. Data were converted to text using the binary mapping tables provided by VicRoads. Each data file contained time and location stamped traffic flow on the through lanes, on-ramps and off-ramps in addition to average vehicle speed on the through lanes with. In order to conduct a comprehensive data analysis, all the data files were combined together in a single database.

The initial traffic data contain detector data from August, 2007 to June, 2010 but traffic data for most of the days of August, 2007 are not available. So traffic data for this month are cleansed and final traffic data begins from September 2007. It is noteworthy to mention that in this study cleansing is done by removing data case from the whole dataset. It is found that sometimes the detectors are not working properly or sometimes there are missing values. For those cases, data cleansing process has been used. The following flow-chart shows the overall process.

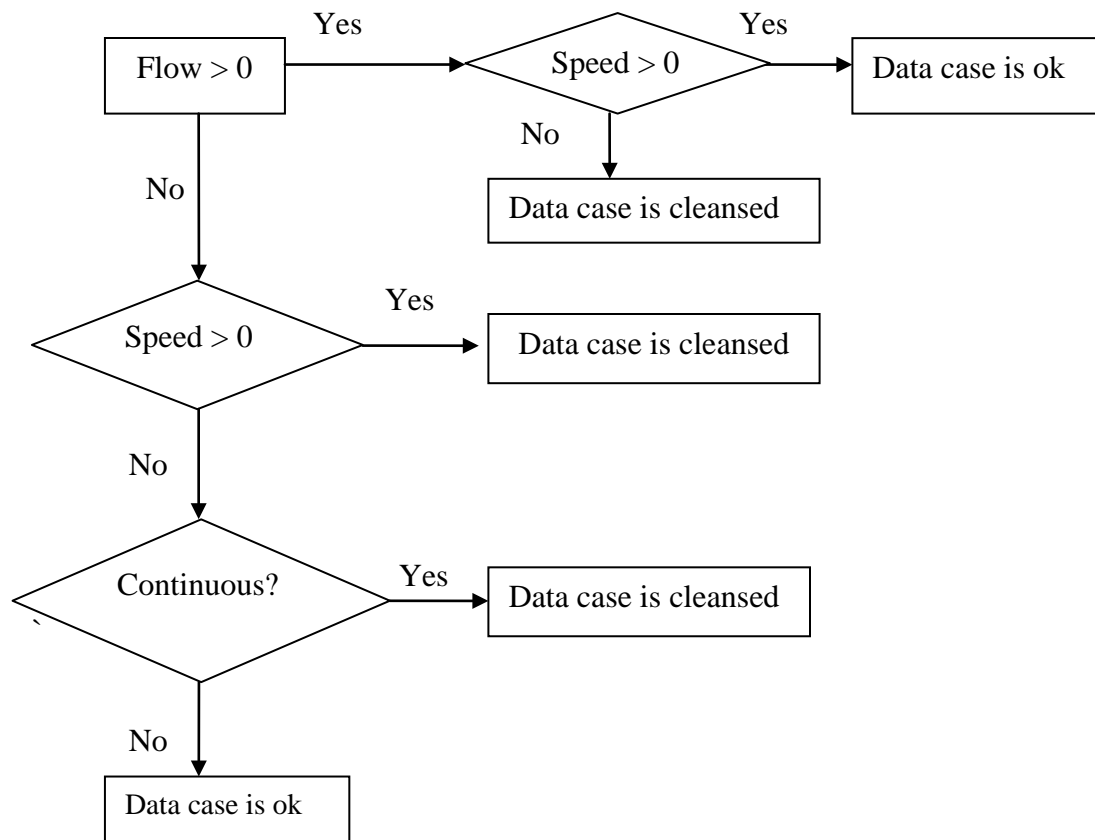


Figure 3.4: Flow-chart of data cleansing process

Final format of the traffic detector data is: date, hour, minute, detector location, traffic flow (through lane, off-ramps and on-ramps) and average speed. Minute by minute traffic data were aggregated to 5 minute intervals in order to smooth the rapid fluctuations providing misleading traffic performance indicators. Aggregation to a 5 minute interval smooths out the fluctuations in the traffic data but at the same time maintains its relevance for its use as an indicator of real-time traffic conditions.

Weather data: Rainfall intensity data by hour are obtained from three weather stations situated near to the test site. Thiessen polygon method (Brassel & Reif, 1979) was applied to find the coverage area corresponding to each weather station and corresponding sections of the Eastern Freeway lying in each weather station's coverage area were identified, It was found that Scorsby Research Institute weather station does not represent the weather condition on any section of the Eastern freeway. Therefore, two other stations, Viewbank (Apsana) and Melbourne regional office are considered. Weather station known as Melbourne regional office represents the weather condition of detector number 1 to 9 in outbound direction and detector number 25 to 33 in inbound direction of the freeway. Remaining sections in each direction are represented by the Viewbank weather stations. These conditions are considered while combining the whole data.

Accident data: The Crash information data for each accident includes accident time, vehicle direction and relative distance from the starting point of the roadway and accident types. This crash information includes all types of accidents including damage only accidents and accidents resulting in injuries or deaths. The total accidents during the data collection period were 138 out of which 62 accidents occurred in outbound direction while the remaining 76 accidents occurred in inbound direction. As no direct locations of the accidents are provided in the accident data relative to the detector locations, spatial concordance between the two datasets was created using detector locations and distance from the starting point of the freeway in Google Maps. Final format of data is: date, hour, minute, location and accident condition.

3.4 Data Fusion

The minute-by-minute traffic data is aggregated to 5 minutes intervals, rainfall intensity data are being kept as hourly data and categorical crash data with specific crash time and locations are used for the analysis. Categorical crash data are set as a binary variable such as: Accident = 'yes', if there was an accident or Accident='no', if there was no accident at the given time and location. Collected dataset indicates only if it is an accident and does not provide information about severity of accidents. Main aim of the study is to develop crash risk prediction model which predicts probability of accident occurrence, so only binary accident data is enough to formulate the model. However, depending on availability of appropriate

data, model functionality can be extended to predict the severity of accidents. This has been listed as a future research direction. To aid the data fusion process, data in each category: traffic, weather and accident are arranged in the same format i.e. date, hour, minute, and detector location. These three data types are used to prepare an aggregated database of 5 minute interval including date, hour, minute, detector location, traffic flow, speed, rainfall intensity during the hour and accident. Data fusion module was developed using programming language Perl. The following flow-chart illustrates the steps in data fusion process.

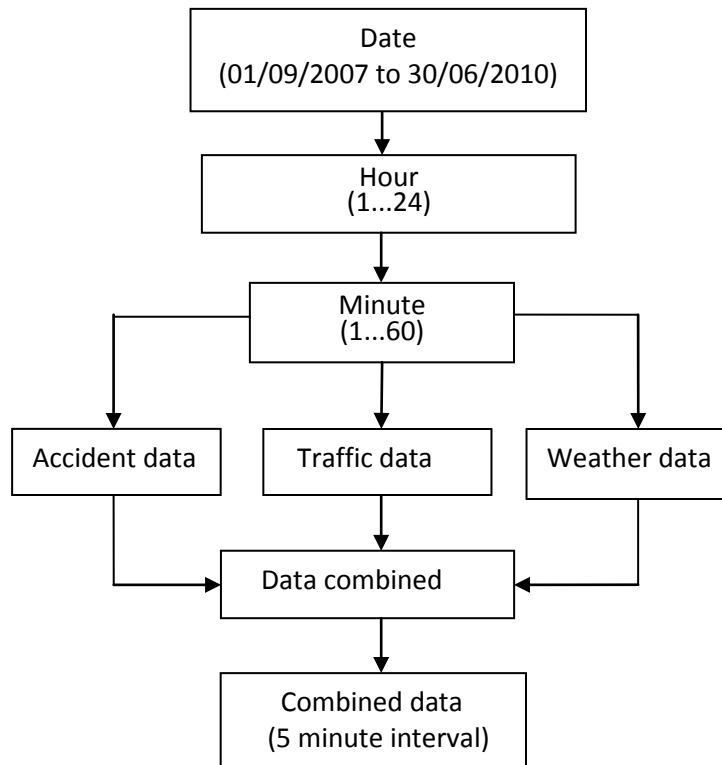


Figure 3.5: Flow-chart of data fusion

3.5 Summary

The database needed for crash risk prediction model requires information from three different source i.e. traffic, weather and accident data. Initially, these datasets were not as organised as needed for model building so they were manipulated, aggregated and fused together to meet the requirements of crash hazard prediction model formulation and validation process.

Chapter 4

MODEL DEVELOPMENT AND VALIDATION

Main objective of this research is to develop a crash hazard prediction model. For this purpose, dataset mentioned in the previous chapter will be used. The developed model will be validated in order to ensure its accuracy and fitness for application. This chapter provides comprehensive description of the two proposed crash hazard prediction models and comparative evaluation in order to select the best model for this study. Two proposed methodologies are classification tree and hazard prediction index.

4.1 Classification Tree Analysis

In order to predict hazard occurrence based on traffic and rainfall data, classification tree analysis is one of the methods adopted in this study. Classification tree models predict categorical dependent variable based on one or more independent variables. A tree like graph which outlines the possible outcomes obtained from particular input variables is used. This method is quite robust, easily understandable and can generate good results despite substantial amount of missing data. Detailed theory about the classification tree analysis can be found in Breiman et al. (1984); Therneau and Atkinson (1997). Here only a brief introduction to the model will be provided.

In general, the dependency of the variables in the classification tree can be shown as

$$Y = f(x_1, x_2, x_3, \dots, x_n) \dots \dots \dots (4.1)$$

Where, Y is the target variable (dependent variable) and $x_1, x_2, x_3, \dots, x_n$ are predictor variables (independent variables).

4.1.1 Building the Classification Tree

In this study, statistical software R (Dalgaard, 2008) is used to develop classification tree by recursive partitioning routine. Recursive partitioning is an exploratory technique to uncover the structure in data. It forms the base for nonparametric regression analysis: Classification and Regression Trees (CART). Model based on this method is built by splitting of dataset into

increasing homogenous subsets until become infeasible to continue (Zhang & Singer, 1999; Higgs & Cummings, 2003).

The process of building the classification tree consists of the following steps:

Step 1: A single predictor variable is chosen for splitting. The predictor variable which can divide total dataset into two parts is selected for first splitting of the tree.

Step 2: The data is separated into two groups based on the chosen split.

Step 3: This process is continued for each subgroup unless it reaches a previously defined minimum size or until any improvement cannot be made (Therneau & Atkinson, 1997).

4.1.2 Splitting Criteria

According to Therneau and Atkinson (1997), if node A is to be split into two branches named left son AL and right son AR, it needs to satisfy the following equation

$$P(AL)I(AL) + P(AR)I(AR) \leq P(A)I(A) \dots \dots \dots (4.2)$$

Where,

$P(A)$ = probability of A

$I(A)$ = risk of A

$P(AL)$ = probability of left son

$I(AL)$ = risk of left son

$P(AR)$ = probability of right son

$I(AR)$ = risk of right son

The split which maximizes the ΔI i.e. decrease the risk, is chosen for building the tree.

$$\Delta I = P(A) I(A) - P(AL) I(AL) - P(AR) I(AR) \dots \dots \dots (4.3)$$

Here,

$$I(A) = \sum p(i|A)L(i, \tau(A)) \dots\dots\dots (4.3.1)$$

where,

$\tau(A)$ = The class method of A, if A were to be taken as a final node.

$p(i|A) = P\{\tau(x) = i \mid x \in A\}$ (for future observations)

$= \pi_i P\{x \in A \mid \tau(x) = i\} / P\{x \in A\}$, thus

$$p(i|A) \approx \pi_i (n_{iA} / n_i) / \sum \pi_i (n_{iA} / n_i) \dots\dots\dots (4.3.2)$$

π_i = prior probabilities of each class; $i = 1, 2, \dots, n$

$L(i, j)$ = Loss matrix for incorrectly classifying an i as a j . $i=1, 2, \dots, C$; $L(i, i) \equiv 0$

n_i, n_A = number of observations in the sample that are in class i , no. of observations in node A.

n_{iA} = number of observations in the sample that are in class i and node A. (Therneau & Atkinson, 1997)

4.1.3 Missing Data Cases

4.1.3.1 Choosing the Split

Recursive partitioning routine (Rpart) can work with missing values within the dataset. Even Rpart takes an observation for analysis which has only one predictor variable with a non-missing value and one target variable. The choosing criterion for any split is to maximise the ΔI .

$$\Delta I = P(A) I(A) - P(AL) I(AL) - P(AR) I(AR) \dots\dots\dots (4.4)$$

Although the first term on the right hand side is same for all variables and splits regardless of missing values, the remaining two terms are somewhat modified for missing data. At first, the

impurity indices of left and right son, $I(AL)$ and $I(AR)$ are calculated for the observations which have non-missing value of the specific independent variables. Then $P(AL)$ and $P(AR)$ are calculated over these observations and are modified to have a sum of $P(A)$. If f be some impurity function, the impurity of a node A is $I(A) = \sum f(p_i A)$; where, $p_i A$ is the proportion of those in A that belong to class i for future samples. If $I(A)=0$ when A is pure, f must be concave with $f(0) = f(1) = 0$. Two candidates for f are the information index, $f(p) = -p \log(p)$ and the Gini index, $f(p) = p(1-p)$. It is found that for the two class problem the measures differ only slightly, and nearly always chooses the same split point (Therneau & Atkinson, 1997).

4.1.3.2 Surrogate Variables

Recursive partitioning routine uses surrogate variables for missing predictor variables in an observation. The surrogate variables can be found by re-applying partitioning algorithm to predict a split for an observation by using other independent variables. For each of the independent variables, a split point and misclassification error is calculated. Then for each split, the surrogates are ranked. The observation with missing split variable is classified by surrogate variables according to the rank until it matches with the value. If surrogate variables cannot be classified as an observation, then "go with majority" rule called blind rule is applied to classify this observation (Therneau & Atkinson, 1997).

4.2 Formulation of Classification Tree Model for this Study

4.2.1 Selection of Model Variables

Dependent variable:

As this classification tree model predicts the hazard occurrence, hazardous situation is considered as the dependent variable where categorical hazard value is either 'Yes (1)' or 'No (0)'. 'Yes (1)' represents hazard condition and 'No (0)' represents non-hazard condition. . It is not possible to measure the probability of accidents in the real world situations. Accident and traffic data are collected from an actual freeway. Accident is a discrete event which has only two likely outcomes i.e. accident or no-accident. Probability of occurrence of these discrete events can be predicted using a mix of discrete i.e. accident data and continuous i.e. traffic data.

Independent variables:

Initially, thirteen independent variables are selected as predictor variables. Traffic flow, vehicle speed and rainfall intensity are considered independent variables in this model. Traffic flow and speed at current time, 5 minutes before, and 10 minutes before at the current detector location and at the nearest upstream detector location are used as independent variables for the modeling purpose. Corresponding rainfall intensity at the same time and location is also used as a predictor variable. To identify variables, q is taken as traffic flow (veh/min), v is taken vehicle speed (km/hr), l is chosen for location and t is chosen for time. Following is a list of variables used in the model:

$q_{l,t}$ = Flow at the current location at current time

$v_{l,t}$ = Speed at the current location at current time

$q_{u,t}$ = Flow at nearest upstream of the current location at current time

$v_{u,t}$ = Speed at upstream of the current location at current time

$q_{l,t-5}$ = Flow at current location at 5 minutes before current time

$v_{l,t-5}$ = Speed at current location at 5 minutes before current time

$q_{u,t-5}$ = Flow at nearest upstream of the current location at 5 minutes before current time

$v_{u,t-5}$ = Speed at nearest upstream of the current location at 5 minutes before current time

$q_{l,t-10}$ = Flow at current location at 10 minutes before current time

$v_{l,t-10}$ = Speed at current location at 10 minutes before current time

$q_{u,t-10}$ = Flow at nearest upstream of the current accident location at 10 minutes before current time

$v_{u,t-10}$ = Speed at nearest upstream of the current location at 10 minutes before current time

R_i = Rainfall intensity

At first, all the variables are input in the software R to build the classification tree but it is found that the following predictors do not have significant contributions in classification tree building:

1. Flow at current location at 5 minutes before current time, $q_{i,t-5}$
2. Flow at current location at 10 minutes before current time, $q_{i,t-10}$
3. Rainfall intensity, R_i

Compared to other predictor variables, the first two variables have a number of missing data points and variation of data which makes these variable less significant than other variables. Although, there were no missing values in rainfall data but it is observed that only approximately 7% of dataset have hazard cases during rainy conditions which are insignificant compared to non-hazard cases and hence makes the use of rainfall intensity in crash hazard prediction trivial. For this reason, rainfall intensity, R_i is not taken into consideration by Rpart routine for tree building.

4.2.2 Building of Classification Tree Model

The classification tree is built by assuming that the minimum number of observations in a terminal node is two and the minimum number of observations in a node for splitting is three. The tree consists of 54 nodes including 28 terminal nodes. The nodes are grown by meeting the criterion of specific range of predictor variables. If any observation meets the criterion, then it moves towards the left side, otherwise it moves to the right. Each node shows the hazard class “yes” or “no” along with the number of progression and non-progressions. The terminal nodes show the hazard and non-hazard occurrence condition by showing 'yes' (hazardous situation) or 'no' (non-hazardous situation). Total classification tree is shown in Figure 4.1.

It is observed from the classification tree that the most important variable for hazard prediction is speed since most of the terminal nodes and split points depend on it. If the predictors are evaluated by time and location, it is seen that the speed and flow at the time and location of the accident occurrence have more contribution to predict crash hazard.

4.2.3 Cross-Validation Result

Table 4.1 shows up to which split the tree can provide accurate results and indicates the best tree for this analysis. The table describes the scaled complexity parameter (cp), relative error ($1-R^2$), cross validation error and cross validation standard deviation value of each split of tree, starting from 0 to 27 splits. Cross validation error is related to the PRESS (Predicted Residual Sums of squares) statistics which is used to provide a summary measure of the fit of a model to a sample of observations and it is calculated as the sum of all resulting errors (Therneau & Atkinson, 1997). The complexity parameter, cp is an advisory parameter and it is specified by the following formula:

$$R_{cp}(T) \equiv R(T) + cp * |T| * R(T_0) \dots\dots\dots(4.5)$$

Where, T_0 is the tree with no splits, $|T|$ is the number of terminal node. A value of $cp = 1$ will result in a tree with no splits. Scaled cp provides direct interpretation in regression model. For example: if any split does not improve the overall R^2 of the model by at least cp, the split is not branched further (Therneau & Atkinson, 1997). The following table depicts that within four splits of the tree, the cross-validation error is minimum while it increases with the addition of number of splits.

Table 4.1: CP table (Errors in each split)

cp	No. of split	Relative error	Cross-validation error	Cross-validated std. deviation
0.347	0	1.000	1.153	0.063
0.057	1	0.653	0.701	0.060
0.036	2	0.596	0.717	0.061
0.016	4	0.524	0.677	0.060
0.013	12	0.395	0.806	0.062
0.011	15	0.355	0.798	0.062
0.010	27	0.169	0.806	0.062

To find out the best classification tree for this analysis 1-SE rule is used which can calculate how many splits of the tree provide good results. In this rule, the split which has minimum value of cross-validation error and minimum value of (cross-validation error + cross-validation standard deviation) is considered as the best tree (Therneau & Atkinson, 1997). According to the 1-SE rule, by four splitting of the tree, the given result is accurate e.g. the prediction of accidents is not over-fitted as this split has minimum value of cross-validation error along with the minimum value of cross validated standard deviation.

The following diagram, cross-validation error vs. size of tree (number of splits +1) depicts the same results as in above table. It is observed that with increasing number of splits of the tree the cross-validation error decreases whereas after four splits (tree size = 5), the cross-validation error increases with the increase in the tree splits. So, instead of the whole tree, it can be pruned after four splits.

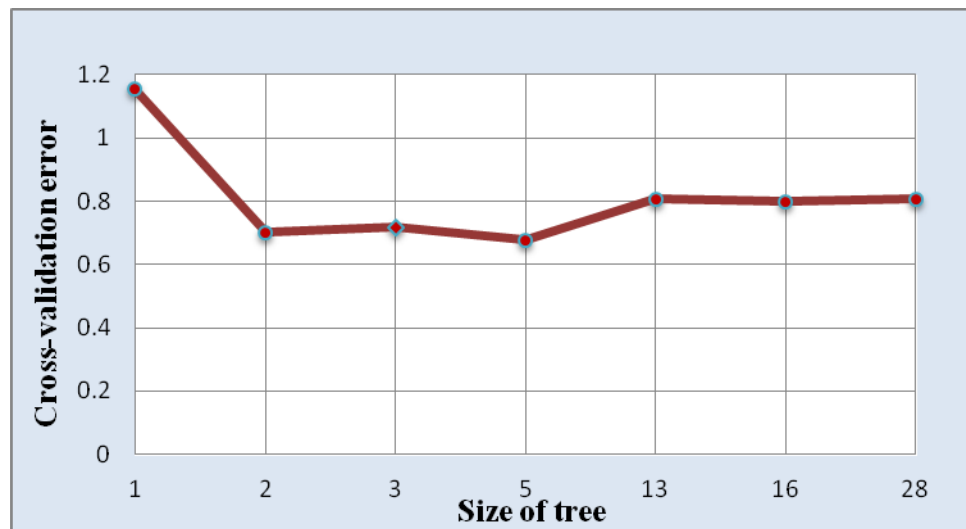


Figure 4.2: Cross-validation error vs. size of the tree

4.2.4 Pruned Classification Tree

If based on the cross-validation results, classification tree is pruned after four splits ($cp=0.016$), it raises another peculiar problem which is explained here. Figure 4.3 shows the pruned tree. The criteria shown for the first right branch of the pruned tree is contradictory since it mentions that if only the speed is less than 85 km/hr, there will be a crash hazard. It predicts hazardous conditions by one criterion and this criterion is set up by one variable

which is speed in this case. This single parameter may be changed due to the factors other than hazard. If the whole classification tree is evaluated, it is found that each branch is terminated before going through at least two criteria. The contradictory situation may have arisen due to limited data. If more cases are included in the dataset, different results may be found after pruning the classification tree. So, due to lack of reliability and accuracy, the pruned classification tree cannot be selected for this study as this will not be able to predict any crash hazard during the high speed conditions. This may also indicate that most of the crash hazard on the test bed site during data collection times have occurred during congested times when the traffic speeds were low.

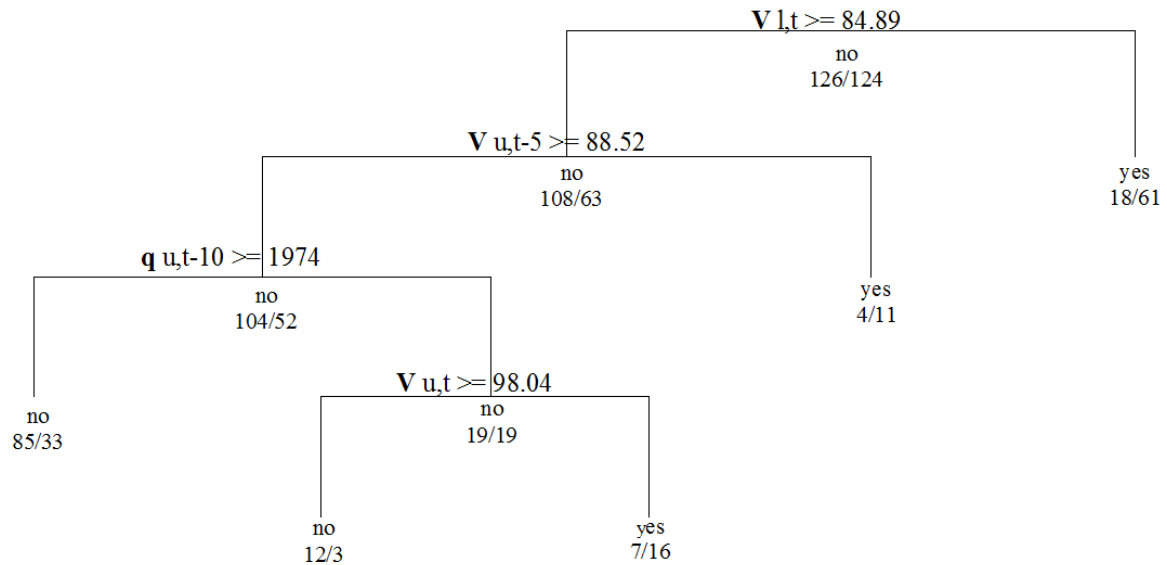


Figure 4.3: Pruned classification tree after cross-validation

4.2.5 Crash Hazard Prediction Model Based on Classification Tree

Since the pruned classification tree provides contradictory prediction, the whole classification tree has been selected as a crash hazard prediction model and then it is validated separately. The following table shows the set of rules based on the classification tree analysis which can be used to classify the traffic situation on the road as hazardous or non-hazardous in terms of crash risk. Table 4.2 helps us to understand how traffic parameters influence hazard and non-hazard situations. The table depicts the traffic condition for two locations i.e. present location ie. where accident occurrence probability needs to be estimated and upstream of accident

occurrence location. For each location, traffic detector data for current time (0 minute before), 5 minute before and 10 minute before is provided. This two dimensional table helps us to understand the evolution of traffic over time and any disturbances in traffic stream which may indicate accident occurrence. Traffic flow and speed value for each of the time are shown in the table. The last column shows the resulting situation on the road for that combination of the traffic parameters i.e. whether it is going to be an accident or not. It is shown that the model develops 28 different combinations of traffic parameters to indicate the hazard situations.

Table 4.2: Accident prediction model (chart view)

No.	Upstream												Location												Rain	Hazard	
	10 min before				5 min before				0 min before				10 min before				5 min before				0 min before						
	Speed		Flow		Speed		Flow		Speed		Flow		Speed		Flow		Speed		Flow		Speed		Flow				
	<	>=	<	>=	<	>=	<	>=	<	>=	<	>=	<	>=	<	>=	<	>=	<	>=	<	>=	<	>=			
1											5130																yes
2											5130			81.69													yes
3											5130			81.69					69.85								no
4											5130			81.69					69.85								yes
5												5130															yes
6						49.69						5130															yes
7							49.69					5130															yes
8							49.69					5130															no
9						88.52		5706																			yes
10						88.52			5706																		no
11				1974			88.52				98.04																no
12				1974			88.52				98.04								96.39								yes
13				1974			88.52				98.04								96.39								yes
14				1638			88.52				98.04								96.39								no
15				1974	1638		88.52				98.04								96.39								yes
16					1974		88.52												92.89								no
17					1974		88.52													92.89							no
18					1974		88.52					98.75	3048							92.89							no
19					1974		88.52					98.75		3048						92.89							yes
20				97.31			88.52				98.75									92.89							no
21		97.31					88.52				98.75		6114							92.89			95	84.89		2184	no
22		97.31					88.52				98.75			6114						92.89			95	84.89		2184	yes
23		97.31					88.52				98.75									92.89			95.97	95		2184	yes
24		97.31					88.52		3840	98.75										92.89						2184	no
25		97.31					88.52	3840		98.75				3288						92.89				95.97	2400	2184	no
26		97.31					88.52	3840		98.75				3288						92.89				95.97		2400	yes
27		97.31						95	3840	98.75		3288								92.89				95.97		2184	no
28		97.31					95	88.52	3840	98.75		3288								92.89				95.97		2184	yes

4.2.6 Model Validation

The dataset used for building the model consists of equal number of hazard and non-hazard cases. All the hazard cases in the raw data are included in the model building dataset while the non-hazard cases are selected for similar weather and traffic conditions at the same location,

and time on a day before the hazardous situation. Before the modeling process, total dataset has been divided into two parts: one part (approximately 82% of the total dataset) is used for model building and another part (about 18% of the total dataset) is used for the validation of the model. Table 4.3 shows the summary of the data used in model building and validation process.

Table 4.3: Summary of the cases used for modeling and validation process

Cases	Model building	Validation
Total	250	53
Hazard	125	17
Non-hazard	125	36

The following histogram shows the overall comparison of the actual data and the predicted values by the model. It is observed that the model predicts few more hazardous situations than the real situation.

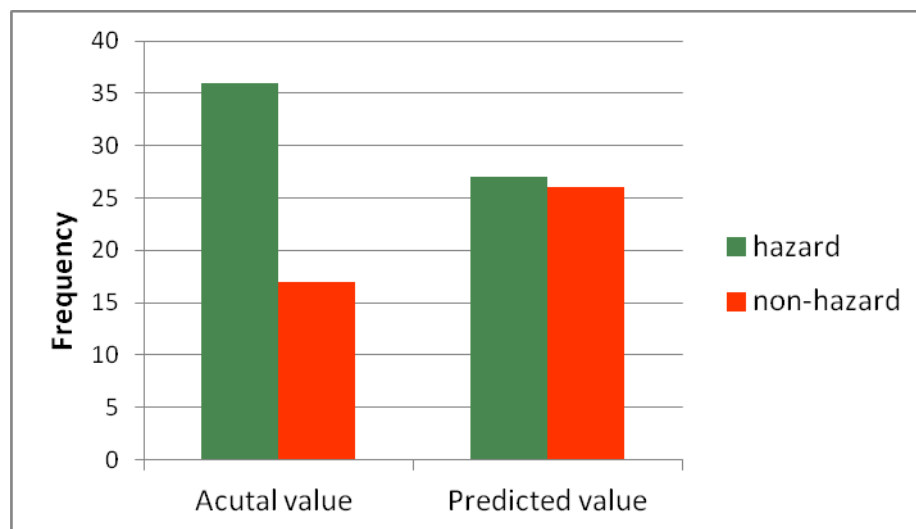


Figure 4.4: Comparative graph of predicted and actual value of the accident and non-accident cases

The following table summarizes the prediction capability of the proposed classification tree model. The measurement of prediction capability is done based on the percentage of error in

predicting the actual situation, the higher the error percentage, the less the prediction accuracy.

Table 4.4: Prediction capability of classification tree

Types of cases	Number of cases	Wrong prediction	% Error
Total	53	13	24.52%
Hazard	17	2	11.76%
Non-hazard	36	11	30.55%

It is observed that the proposed classification tree has an accuracy of 75.48% which is acceptable. The results of Z test also show that this accuracy level is significant. A default model predicting all cases as accidents or non-accidents will gain an accuracy equal to the actual percentage of the corresponding cases as the accuracy level which will be 50% if dataset contains equal number of accident and non-accident cases. To compare the developed model with these hypothetical default models, Z test is conducted.

In this test, 50% accuracy is named as P_1 and 75.48% accuracy is named as P_2 . The total samples for P_1 and P_2 are the number of cases (53 cases) used for validation. The test is one-tailed and the significance level is taken as 0.05. The hypotheses are:

Null hypothesis, $H_0: P_1 \geq P_2$

Alternative hypothesis, $H_a: P_1 < P_2$

The pooled sample proportion,

$$p = \frac{p_1 \times n_1 + p_2 \times n_2}{n_1 + n_2} \dots\dots\dots 4.6$$

The standard error,

$$SE = \sqrt{\{p \times (1-p) \times \left[\left(\frac{1}{n_1} \right) + \left(\frac{1}{n_2} \right) \right] \}} \dots\dots\dots 4.7$$

The Z score,

$$z = \frac{p_1 - p_2}{SE} \dots\dots\dots 4.8$$

In this test, it is found that the value of p , SE and z are 0.6274, 0.0957 and - 2.66 respectively. Since this is one tailed test, the P value is the probability that z score is less than -2.66. By using Z Table, it is found that $P(z < -2.66)$ is 0.00334. Since the P value is less than the significance level, the null hypothesis is rejected and alternative hypothesis is accepted. That means the developed model for this study can perform better than the model of 50% accuracy.

The accuracy level is much higher (about 88%) in accident cases. On the other hand, this model has more prediction error in non-accident cases which is more than 30%. The errors in “Hazard” and “Non-hazard” are so-called “Type I error (false positive)” and “Type II error (false negative)”. In this study, the “Type I errors” are higher than “Type II errors”. It may be because the model predicts the hazardous conditions based on the traffic performance indicators while the driver related characteristics for example, alertness and attention which cannot be observed so easily may have avoided the accidents from happening. It can be said that the model can predict accident cases more correctly than that of non- accident cases.

4.3 Hazard Prediction Index

Although proposed classification tree model can detect the crash hazard significantly, it includes many criteria and boundary conditions. A model which can provide more accurate results with a simple prediction technique will be more useful for real-world real-time applications. Keeping these requirements in mind, a new model is proposed which simplifies the modelling process as well as increases the accuracy.

Hazard Prediction Index is a representative model for predicting the hazardous situations on a roadway by evaluating the traffic performance indicators. The prediction index is built by the difference of traffic parameters with boundary conditions using the same dataset as used in building the classification tree.

4.3.1 Variables Used for Developing the Hazard Prediction Index

The main objective of the model building is to predict the crash hazard based on traffic conditions. So, the independent variables should have the capability to represent the traffic condition. Traffic flow and vehicle speed are such two important variables that can represent traffic condition and were used previously in the classification tree

model. So, traffic performance indicators that were used in building classification tree are also used for developing the hazard prediction index. To identify variables q is taken as traffic flow (veh/min), v is taken vehicle speed (km/hr), l is chosen for location and t is chosen for time.

$q_{l,t}$ = Flow at the current location at current time

$v_{l,t}$ = Speed at the current location at current time

$q_{u,t}$ = Flow at nearest upstream of the current location at current time

$v_{u,t}$ = Speed at upstream of the current location at current time

$q_{l,t-5}$ = Flow at current location at 5 minutes before current time

$v_{l,t-5}$ = Speed at current location at 5 minutes before current time

$q_{u,t-5}$ = Flow at nearest upstream of the current location at 5 minutes before current time

$v_{u,t-5}$ = Speed at nearest upstream of the current location at 5 minutes before current time

$q_{l,t-10}$ = Flow at current location at 10 minutes before current time

$v_{l,t-10}$ = Speed at current location at 10 minutes before current time

$q_{u,t-10}$ = Flow at nearest upstream of the current accident location at 10 minutes before current time

$v_{u,t-10}$ = Speed at nearest upstream of the current location at 10 minutes before current time

4.3.2 Model Formulation

The formulation of the hazard prediction index consists of five steps. The process of this model formulation is described as follows:

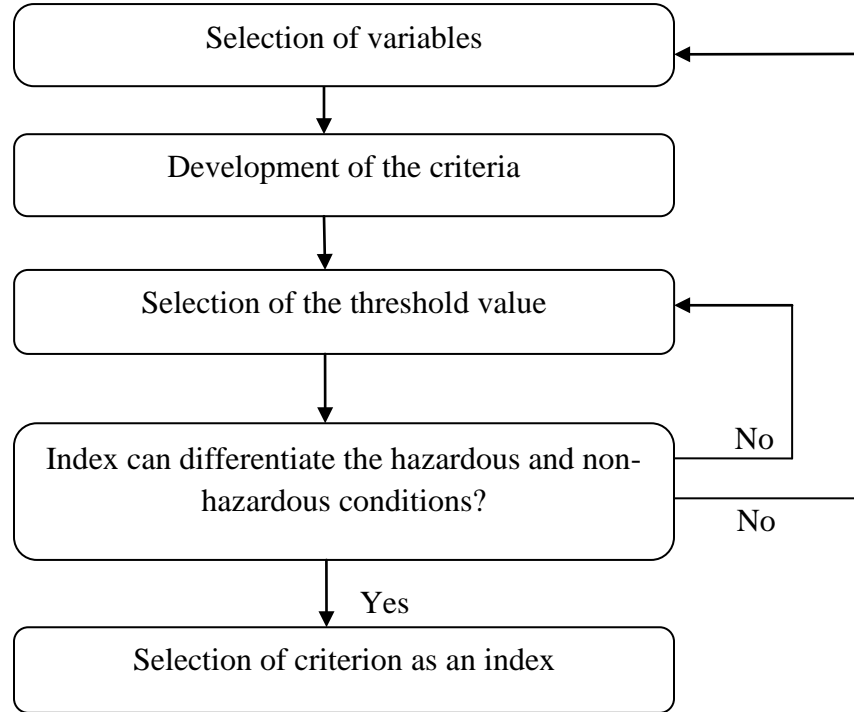


Figure 4.5: Accident prediction index formulation steps

1. ***Selection of variables:*** The first step of model formulation is to select the independent variables from the list of variables mentioned in section 4.3.1. The selection of variables varies from index to index. For each index, three different traffic flow variables and three different speed variables are selected by considering the location and time parameters. Among these three traffic flow and three speed variables, one variable is considered as the ‘controlling variable’ which is exclusive for each index. The controlling variable is used in an index to set the index criteria by measuring the variations of the other selected variables in the index from controlling variable.
2. ***Development of the criteria:*** Every accident probability index consists of four criteria, two of them are for speed and the other two are for traffic flow. The criteria are developed by calculating absolute percentage variation of one variable from the controlling variable. For example, if the controlling variable is $qu, t - 5$ and the other variable is qu, t then the variation percentage would be as following:

$$\{((\text{abs}(q_{u,t} - q_{u,t-5}) * 100 / (q_{u,t-5})))\}$$

Where, abs = absolute value

$q_{u,t}$ = Traffic Flow at the nearest upstream detector location of the current location at current time

$q_{u,t-5}$ = Traffic Flow at the nearest upstream detector location 5 minutes before current time

Similarly, for speed, if the controlling variable is $v_{u,t-5}$ and the other variable is $v_{u,t}$ then the variation percentage would be as following:

$$\{((\text{abs}(v_{u,t} - v_{u,t-5}) * 100 / (v_{u,t-5})))\}$$

Where, abs = absolute value

$v_{u,t}$ = Speed at the nearest upstream detector location of the current location at current time

$v_{u,t-5}$ = Speed at the nearest upstream detector location of the current location 5 minutes before current time

3. ***Initial selection of threshold value:*** Threshold values were used to define the separation between hazardous and non-hazardous traffic conditions. First, two threshold values; one for the traffic flow criteria and other for the speed criteria are selected. For example, in case of the example shown in the previous step; a starting threshold value for flow criteria can be 10 and for speed criteria, it can be 5. So the overall criteria for flow and speed will be as follows:

$$\{((\text{abs}(q_{u,t} - q_{u,t-5}) * 100 / (q_{u,t-5}))) \geq 10 \text{ and}$$

$$\{((\text{abs}(v_{u,t} - v_{u,t-5}) * 100 / (v_{u,t-5}))) \geq 5$$

4. ***Differentiate the hazardous and non-hazardous conditions:*** After the initial selection of the threshold values and four criteria, the focus is to differentiate the hazard and non-hazard conditions. To do so, the criteria are checked in the model building dataset which includes equal number of hazard and non-hazard cases. If

the variables in any case meet the condition of the criteria, it is marked as 1, otherwise 0.

This way, the values against all four criteria are summed up. The process is repeated for each case (accident and non-accident). Later, these values are separated and the average value for hazard and non-hazard cases is calculated. Next, the numerical differences of the average predicted hazard and non-hazard cases are calculated.

5. ***Selection of criterion as an index:*** If the difference of the average values that can separate the hazardous from non-hazardous situations significantly, the criteria are taken as accident prediction index. The main objective is to maximise the difference of the average values of hazardous and non-hazardous cases. The objective function is:

$$\text{maxm of (Avg.value hazard cases — Avg.value nonhazard cases)}$$

In this study, significant values are considered as ratio (average predicted value of non-hazard cases/ average predicted value of hazard cases) that are greater than 50%. If the difference is not significant, then the threshold values are needed to be changed until highest difference can be found. After achieving the highest difference the criterion is considered as the index which can be able to filter the hazardous and non-hazardous traffic conditions.

The threshold value for each of the criteria and combination is selected using an iterative method i.e. trial and error. For each index, the threshold value is kept as same for each of the flow criteria and each of the speed criteria. The trial method for the selection of threshold value starts at 5 and in each steps the value is increased by 5 until a significant value has been found. If several changes of threshold values do not differentiate hazard and non-hazard cases significantly, the criterion is not considered as an index and another criterion is selected for building the index.

4.3.3 Selected Hazard Prediction Index

For this study, seven different hazard prediction indexes are developed which are capable to differentiate hazardous and non-hazardous conditions based on different independent variables. The criterion for each of these indices is mentioned in Table 4.5. Each index includes four criteria and these are divided into two subsets with "and" which means there will only be a hazard if the traffic condition meets the criteria of the two subsets. The hazard prediction index for index 1 can be interpreted as:

It can be said that there will be a hazard if the two criteria are fulfilled at a time

1. The difference between flow at current location at current time and flow at current location 5 minute before current time is 25% of the flow at the current location at current time or the difference between speed at the current location at current time and speed at the current location 5 minute before the current time is 10% of the speed at current location at current time, indicates a hazardous situation.
2. The difference between flow at the current location at current time and flow at upstream location at current time is 25% of the flow of current location at current time or the difference between speed at the current location at current time and speed at the upstream location at current time is 10% of the speed at current location at current time.

Table 4.5: Criteria of Hazard prediction indices

If the any of these two criteria fulfilled, there will be a hazard on the roadway; otherwise there will be no-hazard.

INDEX No.	INDEX CRITERIA
1	$\{(abs(q_{l,t-5} - q_{l,t}) * 100 / (q_{l,t})) \geq 25$ <p style="text-align: center;">or</p> $(abs(v_{l,t-5} - v_{l,t}) * 100 / (v_{l,t})) \geq 10\}$ <p style="text-align: center;">and</p> $\{(abs(q_{l,t} - q_{u,t}) * 100 / (q_{l,t})) \geq 25$ <p style="text-align: center;">or</p> $(abs(v_{l,t} - v_{u,t}) * 100 / (v_{l,t})) \geq 10\}$

2	$\{(abs(q_{l,t-5} - q_{u,t-10}) * 100 / (q_{u,t-10})) \geq 25$ <p style="text-align: center;">or</p> $abs(v_{l,t-5} - v_{u,t-10}) * 100 / (v_{u,t-10})) \geq 10\}$ <p style="text-align: center;">and</p> $\{(abs(q_{l,t} - q_{u,t-10}) * 100 / (q_{u,t-10})) \geq 25$ <p style="text-align: center;">or</p> $(abs(v_{l,t} - v_{u,t-10}) * 100 / (v_{u,t-10})) \geq 10\}$
3	$\{(abs(q_{l,t-10} - q_{l,t-5}) * 100 / (q_{l,t-5})) \geq 20$ <p style="text-align: center;">or</p> $(abs(v_{l,t-10} - v_{l,t-5}) * 100 / (v_{l,t-5})) \geq 5\}$ <p style="text-align: center;">and</p> $\{(abs(q_{l,t} - q_{l,t-5}) * 100 / (q_{l,t-5})) \geq 20$ <p style="text-align: center;">or</p> $(abs(v_{l,t} - v_{l,t-5}) * 100 / (v_{l,t-5})) \geq 5\}$
4	$\{(abs(q_{u,t-10} - q_{u,t}) * 100 / (q_{u,t})) \geq 20$ <p style="text-align: center;">or</p> $(abs(v_{u,t-10} - v_{u,t}) * 100 / (v_{u,t})) \geq 10\}$ <p style="text-align: center;">and</p> $\{(abs(q_{l,t-5} - q_{u,t}) * 100 / (q_{u,t})) \geq 20$ <p style="text-align: center;">or</p> $(abs(v_{l,t-5} - v_{u,t}) * 100 / (v_{u,t})) \geq 10\}$
5	$\{(abs(q_{l,t-5} - q_{l,t-10}) * 100 / (q_{l,t-10})) \geq 20$ <p style="text-align: center;">or</p> $(abs(v_{l,t-5} - v_{l,t-10}) * 100 / (v_{l,t-10})) \geq 5\}$ <p style="text-align: center;">and</p> $\{(abs(q_{l,t} - q_{l,t-10}) * 100 / (q_{l,t-10})) \geq 20$ <p style="text-align: center;">or</p> $(abs(v_{l,t} - v_{l,t-10}) * 100 / (v_{l,t-10})) \geq 5\}$

6	$\{(abs(q_{u,t} - q_{u,t-5}) * 100 / (q_{u,t-5})) \geq 25$ <p style="text-align: center;">or</p> $(abs(v_{u,t} - v_{u,t-5}) * 100 / (v_{u,t-5})) \geq 5\}$ <p style="text-align: center;">and</p> $\{(abs(q_{l,t-5} - q_{u,t-5}) * 100 / (q_{u,t-5})) \geq 25$ <p style="text-align: center;">or</p> $(abs(v_{l,t-5} - v_{u,t-5}) * 100 / (v_{u,t-5})) \geq 5\}$
7	$\{(abs(q_{u,t-5} - q_{u,t-10}) * 100 / (q_{u,t-10})) \geq 15$ <p style="text-align: center;">or</p> $(abs(v_{u,t-5} - v_{u,t-10}) * 100 / (v_{u,t-10})) \geq 5\}$ <p style="text-align: center;">and</p> $\{(abs(q_{l,t-10} - q_{l,t-10}) * 100 / (q_{l,t-10})) \geq 25$ <p style="text-align: center;">or</p> $(abs(v_{l,t-10} - q_{l,t-10}) * 100 / (q_{l,t-10})) \geq 5\}$

4.3.4 Validation of the Indices

Total dataset was divided into two subsets one each for building and validation of the hazard prediction index. Model building subset consists of 80% of the total data and it is used for index formulation while model validation subset, consists of 20% of the total data and is used for validation of the index.

Table 4.6 compares the prediction capability of each index. As predicting hazardous situations leading to accidents is the main objective of the model building, strength of the each of the indices is evaluated both for total cases (accident and non-accident cases) and accident-only cases. The measurement of prediction capability of the indices is done with the percentage of error in predicting the actual situation, the higher the error percentage, the less the capability.

Table 4.6: Prediction capability of the accident prediction indices

Index Number	Total cases (hazard and non-hazard)	Total wrong prediction	% Error in total cases
1	53	17	32.07%
2	53	14	26.41%
3	53	14	26.41%
4	53	13	24.52%
5	53	10	18.86%
6	53	10	18.86%
7	53	16	30.18%

It is evident from the above table that the differences in percentage of prediction error of indices are large. And it is observed that index 5 and index 6 are well validated compared to others since these indices have the least prediction error which less than 20%.

4.3.5 Selection of the Best Accident Prediction Index

In the context of overall prediction capability, index 5 and index 6 have the least prediction error whereas index 1 and index 7 have quite higher value than that of index 2, index 3 and index 4. So, considering the overall prediction capability, index 5 or index 6 can be the best choice. But for hazard-only cases, the scenario is quite different. In comparing the percentage error in prediction of hazard-only cases, it is found that Index 5 and Index 6 have percentage error of 52.94% and 35.3 % respectively. So, it can be said that for hazard-only cases, index 6 can be the best predictor of the hazardous conditions on the roadway. After comparing both types of error, it is decided that index 6 is the best crash prediction index.

4.4 Comparative Analysis of Classification Tree and Accident Prediction Model

Previous sections of this chapter have shown overall description and validation process of the two crash hazard prediction models: classification tree model and hazard prediction index. Although the same dataset is used for both models, it has been observed that Hazard prediction

index has greater accuracy than Classification tree model. Table 4.7 depicts the overall comparison of these two selected model.

Table 4.7: Comparison of prediction capability of two selected models

Model	Percentage of False Prediction	Percentage of Accuracy
Classification tree model	24.52%	75.48%
Hazard prediction index	18.86%	81.14%

It is observed that the proposed classification tree has an accuracy of 75.48% and hazard prediction index has 81.14% accuracy. To ensure that hazard prediction index has better performance, hypothesis test known as Z test is conducted.

In this test, the percentage gained by hazard prediction index is named as P_1 and the percentage gained by classification tree model is named as P_2 . The total samples for P_1 and P_2 are the number of cases (53 cases) used for validation. The test is one-tailed and the significance level is taken as 0.05. The hypotheses are:

Null hypothesis, $H_0: P_1 \geq P_2$

Alternative hypothesis, $H_a: P_1 < P_2$

By using the equation 4.6, 4.7, 4.8, it is found that the value of p, SE and z are 0.7831, 0.0816 and 0.693 respectively.

Since this is one tailed test, the P value is the probability that z score is less than 0.693. By using Z Table, it is found that $P(z < 0.693)$ is 0.773. Since the P value is greater than the significance level, the null hypothesis cannot be rejected. That means the hazard prediction index can perform better than the classification tree model.

In addition, the selected hazard prediction index is based on six predictor variables whereas the classification tree model has ten independent variables based on traffic flow and speed at different times and locations. Also, the proposed classification tree model consists of twenty eight different criteria to predict crash hazard conditions whereas the selected index specifies

only two criteria. So, it can be said that hazard prediction index is easy to understand and implement for real-time real world applications.

Furthermore, the criteria in the classification tree are developed by the range of values of independent variables and in some cases these values are in decimal forms. On the other side, the hazard prediction index is formulated with certain differences of predictor values. Thus, it is less liable to changes than that of classification tree since a small decimal change in classification tree predictor's values can wrongly indicate a hazardous condition as non-hazard.

4.5 Selected Crash Risk Probability Model for the Study

From the overall comparison of the two previously proposed prediction models, it can be concluded that hazard prediction index model is superior to classification tree model in the context of accuracy, complexity and vulnerability. Thus, for this study, hazard prediction index model is selected as crash risk probability model.

The Crash risk probability model is as follows where q is traffic flow (veh/min), v is vehicle speed (km/hr), l is indicates location and t indicates time.

If the any of the following two criteria are fulfilled, a crash hazard situation is indicated on the roadway.

$$\begin{aligned} & \{((\text{abs}(q_{u,t} - q_{u,t-5}) * 100 / (q_{u,t-5})) \geq 25 \\ & \quad \text{or} \\ & \quad (\text{abs}(v_{u,t} - v_{u,t-5}) * 100 / (v_{u,t-5})) \geq 5\} \\ & \quad \text{and} \\ & \{((\text{abs}(q_{l,t-5} - q_{u,t-5}) * 100 / (q_{u,t-5})) \geq 25 \\ & \quad \text{or} \\ & \quad (\text{abs}(v_{l,t-5} - v_{u,t-5}) * 100 / (v_{u,t-5})) \geq 5\} \end{aligned}$$

Where,

$q_{u,t}$ = Flow at nearest upstream of the current location at current time

$v_{u,t}$ = Speed at upstream of the current location at current time

$q_{l,t-5}$ = Flow at current location at 5 minutes before current time

$v_{l,t-5}$ = Speed at current location at 5 minutes before current time

$q_{u,t-5}$ = Flow at nearest upstream of the current location at 5 minutes before current time

$v_{u,t-5}$ = Speed at nearest upstream of the current location at 5 minutes before current time

The selected criteria mean that there will be a hazard if the two criteria are fulfilled at a time:

1. The difference between flow at upstream at current time and flow of upstream at 5 minute before current time is 25% of the flow at upstream at 5 minute before current time or the difference between speed at upstream at current time and speed at upstream at 5 minute before current time is 5% of the speed at upstream at 5 minute before current time.
2. The difference between flow at the current location at 5 minute before current time and flow at upstream at 5 minute before current time is 25% of the flow at upstream at 5 minute before current time or the difference between speed at the current location at 5 minute before current time and speed at upstream at 5 minute before current time is 5% of the speed at upstream at 5 minute before current time.

It can be said that there are four combinations of traffic parameters found from the selected crash risk probability model and each of these combinations leads to a hazardous condition.

These are:

1. $(\text{abs}(q_{u,t} - q_{u,t-5}) * 100 / (q_{u,t-5})) \geq 25$ and $(\text{abs}(q_{l,t-5} - q_{u,t-5}) * 100 / (q_{u,t-5})) \geq 25$
2. $(\text{abs}(q_{u,t} - q_{u,t-5}) * 100 / (q_{u,t-5})) \geq 25$ and $(\text{abs}(v_{l,t-5} - v_{u,t-5}) * 100 / (v_{u,t-5})) \geq 5$

$$3. (\text{abs}(v_{u,t} - v_{u,t-5}) * 100 / (v_{u,t-5})) \geq 5 \text{ and } (\text{abs}(q_{l,t-5} - q_{u,t-5}) * 100 / (q_{u,t-5})) \geq 25$$

$$4. (\text{abs}(v_{u,t} - v_{u,t-5}) * 100 / (v_{u,t-5})) \geq 5 \text{ and } (\text{abs}(v_{l,t-5} - v_{u,t-5}) * 100 / (v_{u,t-5})) \geq 5$$

4.6 Summary

Two crash hazard risk probability model have been analysed and validated with real data and have shown an acceptable percentage of accuracy. Among these two models, the best one, hazard prediction index is selected. This model will be used to aid the process of road hazard mitigation schemes. It will be used to detect the hazardous traffic conditions and based on the hazard conditions a mitigation strategy will be implemented.

Chapter 5

SIMULATION STUDY

Simulation is an effective tool to investigate the efficacy of the proposed model in devising suitable mitigation schemes to reduce the hazardous conditions on the road before implementing it in the field. This chapter illustrates an overview of the simulation process including calibration and validation of the simulation model.

5.1 Simulation Features

A microscopic traffic simulation package known as AIMSUN (Advanced Interactive Microscopic Simulator for Urban and Non-Urban Networks) (Barceló & Casas, 2002) version 6.0 is used in this study. It can emulate the real world scenario with detailed road network and traffic elements (TSS-Transport Simulation Systems, 1997). This package is capable of checking the traffic control strategies and traffic management policies based on traditional or Intelligent Transport Systems.

During the run time of the simulation, the states of some elements for example detectors and vehicles are updated continuously whereas the states of other elements; traffic signals, ramp metering, speed limit etc. can be fixed or changed discretely throughout the simulation time. The input data can be divided into two types: scenario and simulation parameters. The scenario is formed by four data types: network layout, traffic demand data, traffic control system and public transport plans (if applicable). On the other hand, simulation parameters consist of fixed values of simulation time, warm-up period, statistics intervals, car following model parameters etc. The outputs from AIMSUN include graphical representation of the traffic network performance (in 2D AND 3D), statistical data (flow, speed, travel time, delays, stop time etc.) and simulated detector data (counts, speed, occupancy etc.) (TSS - Transport Simulation Systems, 1997).

5.2 Input Data for Simulation

5.2.1 Network Layout

Traffic network consists of different interconnecting sections which may not have similar features. For this study, 18 kilometre long route of both inbound and outbound direction of Eastern Freeway with details of the number of lanes, width, speed limit for each section, and details of on-ramps, off-ramps for both direction (inbound and outbound) are drawn to build the network. This is done with the aid of the GIS file (base map), Google Earth and location data.

5.2.2 Traffic Demand Data

In AIMSUN package, traffic demand data can be input in two ways: one is by the traffic flows at the sections and the other is by O/D matrix (TSS-Transport Simulation Systems, 1997). In this study, hourly sliced O/D matrix based demand data are used. All the on-ramps and the starting entry point are taken as origins while all the off-ramps and final exit point are taken as destinations. As the selected freeway is a linear road network, route choice is not needed for calibration of demand data. Typical origin-destination matrices showing the possible OD patterns for the selected freeway are shown in Table 5.1 and Table 5.2 where O is used for origin and D is used for destination.

5.2.2.1 OD Matrix for Outbound Direction

In outbound direction, O_1 and D_8 are the entry and exit point of the freeway whereas the other origins ($O_2, O_3, O_4, O_5, O_6, O_7$) and destinations ($D_1, D_2, D_3, D_4, D_5, D_6, D_7$) are the on-ramps and off-ramps of this roadway.

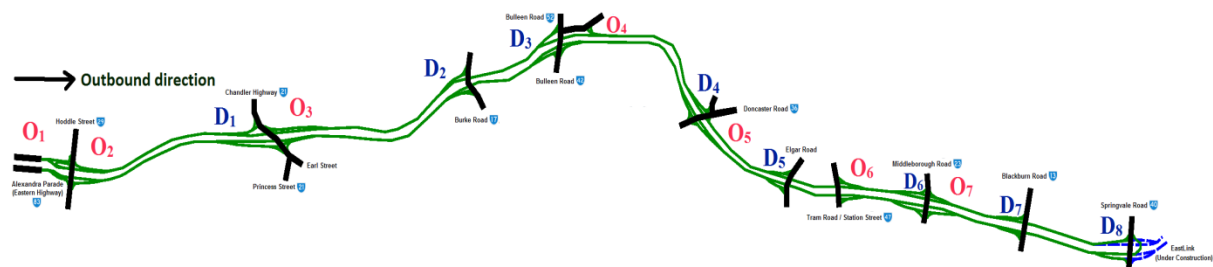


Figure 5.1: Origins and destinations for outbound direction of Eastern freeway

Table 5.1 shows the framework of OD table for outbound direction of the freeway. It is seen that few cells in the OD table are empty. This is due to the fact that some on-ramps are situated after few destinations. Thus there is no traffic flow from these origins to destinations.

Table 5.1: Framework for OD matrix in outbound direction

Destination \ Origin	D ₁	D ₂	D ₃	D ₄	D ₅	D ₆	D ₇	D ₈
O ₁	O ₁ D ₁	O ₁ D ₂	O ₁ D ₃	O ₁ D ₄	O ₁ D ₅	O ₁ D ₆	O ₁ D ₇	O ₁ D ₈
O ₂	O ₂ D ₁	O ₂ D ₂	O ₂ D ₃	O ₂ D ₄	O ₂ D ₅	O ₂ D ₆	O ₂ D ₇	O ₂ D ₈
O ₃		O ₃ D ₂	O ₃ D ₃	O ₃ D ₄	O ₃ D ₅	O ₃ D ₆	O ₃ D ₇	O ₃ D ₈
O ₄				O ₄ D ₄	O ₄ D ₅	O ₄ D ₆	O ₄ D ₇	O ₄ D ₈
O ₅					O ₅ D ₅	O ₅ D ₆	O ₅ D ₇	O ₅ D ₈
O ₆						O ₆ D ₆	O ₆ D ₇	O ₆ D ₈
O ₇							O ₇ D ₇	O ₇ D ₈

5.2.2.2 OD Matrix for Inbound Direction

In this direction, O₁ and D₇ are the entry and exit points of the freeway whereas the other origins (O₂, O₃, O₄, O₅, O₆, O₇, O₈, O₉) and destinations (D₁, D₂, D₃, D₄, D₅, D₆) are the on-ramps and off-ramps respectively.

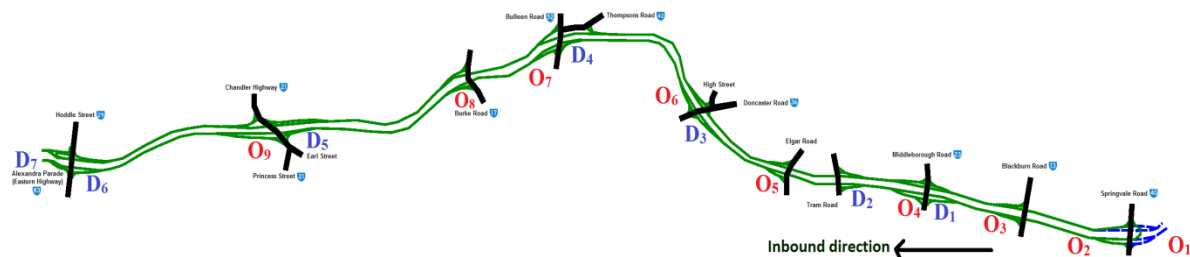


Figure 5.2: Origins and destinations for inbound direction of Eastern freeway

Table 5.2 shows the framework of OD table for inbound direction. It is seen that some cells in the OD table are empty. This is because of the fact that some on-ramps are situated after few destinations. Thus there is no traffic flow from these origins to destinations.

Table 5.2: Framework for OD matrix in inbound direction

Destination \ Origin	D₁	D₂	D₃	D₄	D₅	D₆	D₇
O₁	O ₁ D ₁	O ₁ D ₂	O ₁ D ₃	O ₁ D ₄	O ₁ D ₅	O ₁ D ₆	O ₁ D ₇
O₂	O ₂ D ₁	O ₂ D ₂	O ₂ D ₃	O ₂ D ₄	O ₂ D ₅	O ₂ D ₆	O ₂ D ₇
O₃	O ₃ D ₁	O ₃ D ₂	O ₃ D ₃	O ₃ D ₄	O ₃ D ₅	O ₃ D ₆	O ₃ D ₇
O₄		O ₄ D ₂	O ₄ D ₃	O ₄ D ₄	O ₄ D ₅	O ₄ D ₆	O ₄ D ₇
O₅			O ₅ D ₃	O ₅ D ₄	O ₅ D ₅	O ₅ D ₆	O ₅ D ₇
O₆					O ₆ D ₅	O ₆ D ₆	O ₆ D ₇
O₇					O ₇ D ₅	O ₇ D ₆	O ₇ D ₇
O₈					O ₈ D ₅	O ₈ D ₆	O ₈ D ₇
O₉						O ₉ D ₆	O ₉ D ₇

5.2.3 Vehicle Configuration

Traffic data provided by VicRoads include traffic flow rate regardless of vehicle types. Also, there is no available literature to get the approximate percentage of trucks on this freeway. so, car is chosen as the only vehicle type although this may skew the results.

5.3 Calibration Procedure for Simulation Output

Calibration is an iterative process that consists of changing model parameters and comparing model outputs with a set of real data until a predefined level of agreement between the two data sets is achieved (TSS-Transport Simulation Systems, 1997). In this study, calibration process is executed by estimating the origin-destination matrix, calculating vehicle speed and selection of the suitable parameters for calibration which is used to check the accuracy of the simulation outputs. Figure 5.3 depicts the calibration framework. To be a calibrated model,

the difference of simulated and actual data should be below a certain predefined threshold. In this study, the threshold is taken as 10% variation of simulated speed from actual speed.

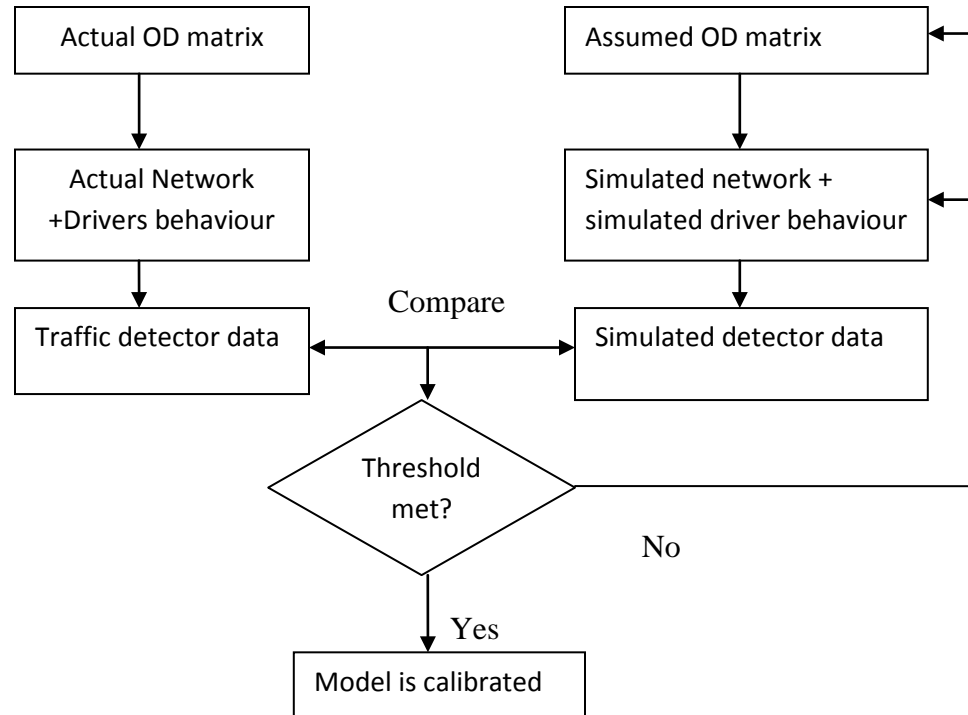


Figure 5.3: Framework for calibration process

5.3.1 Selected Parameter for Calibration

Traffic flow may not be a better choice as a parameter for calibration because the same flow value can be reached in congested and free flow conditions. This can be understood from the fundamental diagram of traffic flow. For this reason, traffic counts are not able to indicate congestion and non-congestion situation (Bevrani et al., 2011). However, speed values can differentiate the two situations properly. So, speed rather than traffic count is selected as the target parameter for calibration.

Calibration of the simulation data is conducted by comparing the actual speed data collected from traffic detectors with the simulated speed data for the same time and location. It is also worthy to mention that traffic flow values are not totally ignored as these are used to calibrate the OD matrix for the traffic simulation.

5.3.2 Origin-Destination Flow Estimation

As the dataset used for this study only has traffic detector flows and speeds and origin-destination flows are not known, an OD matrix needs to be estimated. In this case, an iterative process based on trial and error method is implemented in order to make a reliable hourly OD matrix. This is done by applying "Populate OD table strategy". To make the OD table, this strategy uses hourly traffic flow data and OD flow patterns as shown in Table 5.1 and Table 5.2 for both inbound and outbound direction of the freeway. From the traffic flow data, the total numbers of vehicles entering and departing from each on-and off-ramp per hour are found. Based on the total flow, the traffic flow for each OD pair is calculated using iterative trial and error method until the total flow for each on- and off-ramp is matched with the traffic detector data. It is worthy to mention that due to the presence of detectors on the on-ramps and off-ramps, total flows from each on-ramp and off-ramp are known in 5 min intervals. The unknown parameter is how to distribute this total flow from origin to each possible destination or from destination to each possible origin. The overall process to achieve this goal is illustrated in Figure 5.4.

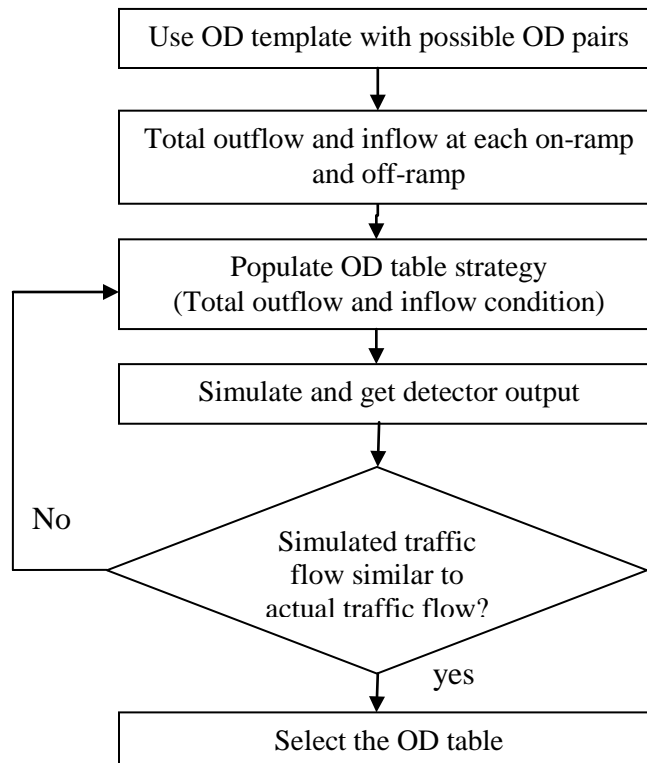


Figure 5.4: OD matrix estimation process

Traffic flow rate of the detectors can be related with the OD's flow rate. When the cell values in OD table are changed, there is a change in detector flow values. So, the iteration process will continue until difference reasonable agreement between these simulated hourly flow values and real traffic data is obtained. For this study, less than 5% variation in traffic flow rate is considered as reasonable. Table 5.3 shows an overview of the building of an OD matrix table. The following OD table represents the OD matrix of outbound direction on 16/11/2009 during 4:00 pm to 5:00 pm. As this OD table meets the reasonable agreement of 5% variation of traffic flow between simulated and actual situation, this table is selected for analysis.

Table 5.3: Selected OD table

Origin/Destination	D1	D2	D3	D4	D5	D6	D7	D8		Total (simulated)	Total (actual)
O1	200	405	732	200	217	168	239	3584		5745	5745
O2	72	100	50	50	50	50	50	115		537	537
O3		200	500	200	100	200	50	397		1647	1647
O4				150	100	150	50	510		960	960
O5					100	100	50	329		579	579
O6						250	350	663		1263	1263
O7							100	106		206	206
Total (simulated)	272	705	1282	600	567	918	889	5704			
Total (actual)	272	705	1282	600	567	918	889	5516			
Detector	Flow (actual)		Flow (simulated)			Difference					
d1	5745		5745			0					
d3	6258		6282			-24					
d6	5924		6010			-86					
d8	7497		7657			-160					
d12	6688		6952			-264					
d15	5393		5670			-277					
d17	6345		6630			-285					
d19	5738		6030			-292					
d21	6331		6609			-278					
d22	5744		6042			-298					
d25	7110		7305			-195					
d26	6138		6387			-249					
d27	3965		4177			-212					
d29	5575		5704			-129					

5.3.3 Calculating the Speed of a Vehicle in a Section

Speed is one of the key parameters for traffic simulation calibration in this study. It is necessary to calibrate the simulated speed to get a reliable simulation output. Three parameters are used in AIMSUN6.0 to calculate the maximum desired speed of a vehicle: two are related to the vehicle and one to the section (TSS-Transport Simulation Systems, 1997):

- Maximum desired speed of the vehicle $i = v_{\max}(i)$
- Speed acceptance of the vehicle $i = \theta(i)$
- Speed limit of the section $s = S_{\text{limit}}(s)$

The parameter $\theta(i)$ can be interpreted as the "level of goodness" of the drivers or the degree of acceptance of speed limits. $\theta(i) \geq 1$ means that the vehicle i will take as maximum speed for a section a value greater than the speed limit, while $\theta(i) \leq 1$ means that the vehicle i will use a lower speed limit (TSS-Transport Simulation Systems, 1997).

The speed limit for a vehicle i on a section s , $S_{\text{limit}}(i,s)$, is calculated as:

$$S_{\text{limit}}(i,s) = S_{\text{limit}}(s) * \theta(i) \dots\dots\dots(5.1)$$

Then the maximum desired speed of vehicle i on a section s , $v_{\max}(i,s)$ is calculated as:

$$v_{\max}(i,s) = \text{minimum} [S_{\text{limit}}(i,s), v_{\max}(i)] \dots\dots\dots(5.2)$$

For the study, $v_{\max}(i)$ and $\theta(i)$ are taken as constant regardless of variable i as only one type of vehicle "car" is considered throughout the simulation process. For all the simulation, the values of the three parameters are taken as:

$$v_{\max}(i) = 110 \text{ km/hr}$$

$$\theta(i) = 0.95$$

$$S_{\text{limit}}(s) = 100 \text{ km/hr}$$

The value of $v_{\max}(i)$ and $S_{\text{limit}}(s)$ are taken as same as used in the freeways in Victoria, Australia. And the value of $\theta(i)$ is selected by evaluating the actual speed dataset where it is found that most of the speed in normal traffic flow condition have speeds in the range of 95 km/hr to 100 km/hr.

So, the speed limit for a vehicle i on a section s , $S_{\text{limit}}(i,s) = 95 \text{ km/hr}$

And the maximum desired speed of vehicle i on a section s ,

$$v_{\max}(i,s) = \text{minimum} [95, 110] = 95 \text{ km/hr}$$

The calculated value is approximate to the actual data. Thus the maximum desired speed considered for simulation can represent the actual situation accurately.

Figure 5.5 shows the parameters used in the simulation process. These parameters are set by iterative process i.e. trial and error method. After each trial, the simulation output is checked for accuracy. Finally, the following parameters are found as the best representative of driving behaviour.

The screenshot shows a software window titled "Vehicle Type: 8, Name: car". It contains several tabs: "Main", "Classes", "Articulated", "2D Shapes", "3D Shapes", "Experiment Defaults", "Fuel", "Pollutants", and "Attributes". The "Main" tab is active, displaying a table of parameters and their values.

Name	Mean	Deviation	Min	Max	Units
Length	4	0.5	3.4	4.6	meters
Width	2	0	2	2	meters
Max Desired Speed	110	10	100	110	km/h
Max Acceleration	3	0.2	2.6	3.4	m/s ²
Normal Deceleration	4	0.25	3.5	4.5	m/s ²
Max Deceleration	6	0.5	5	7	m/s ²
Speed Acceptance	0.95	0.1	0.94	0.96	
Min Distance Veh	2	0.1	1.9	2.5	meters
Give Way Time	10	2.5	5	15	Secs
Guidance Acceptance	75	10	65	90	%
Sensitivity Factor	1	0	1	1	
Minimum Headway	0	0	0	0	Secs

Below the table, there are several input fields for specific parameters:

- After overtaking stay on fast lane: %
- Undertaking cases: %
- Imprudent Lane Changing cases: %
- Sensitivity for Imprudent Lane Changing:
- Equipped Vehicles: %
- Cruising Tolerance: m/s²
- PCUs:

At the bottom right, there are "OK" and "Cancel" buttons.

Figure 5.5: Parameters used for selected vehicle types

5.3.4 Calibration Results

For calibration purposes, traffic conditions on a normal day for each direction are selected. The selected days are:

- Outbound direction: date: 29/10/2009, time: 2 p.m. to 3:00 p.m.
- Inbound direction: date: 23/05/2009, time: 4 p.m. to 5:30 p.m.

The simulated speed, actual speed and difference between these values for all the sections of the roadway are presented in contour plot. In the contour plot, speed ranges are kept as 20

km/hr in case of actual and simulated speed illustration whereas ranges are taken as 10 km/hr for the illustration of speed difference. In this study, the simulation model is considered as calibrated if the difference between actual speed and simulated speed is less than 10 km/hr in most of the cases.

It is observed that the simulated speed values closely match the actual speed values. Apart from few outliers, most of the simulated speeds at all detector locations during simulation time are within the acceptable range of 10% difference from actual speeds. Following figures show the results of the calibration process:

Case 1: Outbound direction (29/10/2009)

It is observed from the Figure 5.6 (c) that the difference in simulated speed and actual data is less than 10 km/hr for all of the sections. So, the model is considered as a well calibrated model.

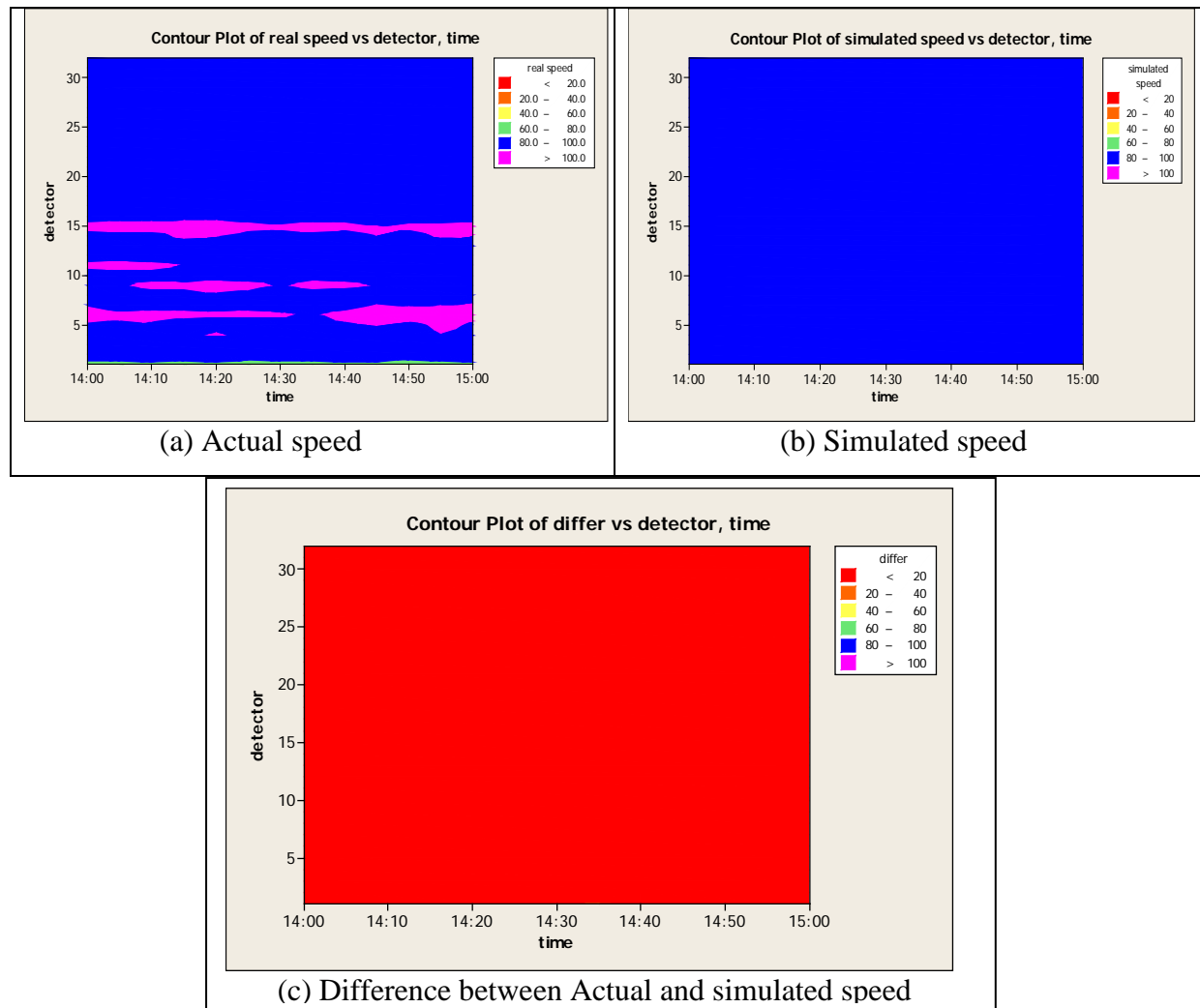


Figure 5.6: Calibration result for outbound sections on 29/10/2009

Case 2: Inbound direction (23/05/2009)

It is observed that the difference in simulated speed and actual speed is less than 10 km/hr for most of the sections. However, in few sections, the difference is in the range of 10 km/hr to 20 km/hr. So, it can be said that the model is a well calibrated model. Figure 5.7 represents the overall results.

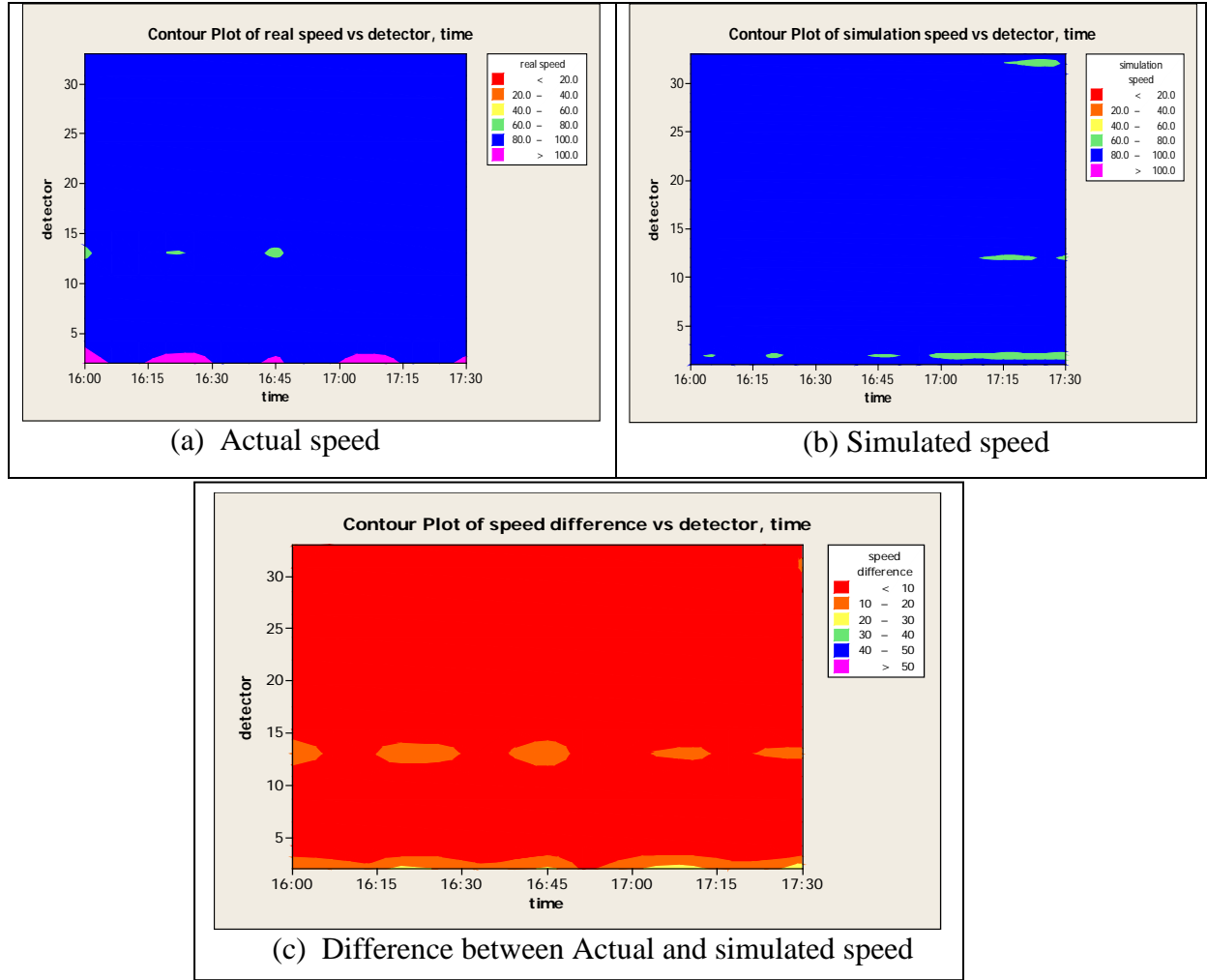


Figure 5.7: Calibration result for inbound sections on 23/05/2009

5.4 Validation Process for Simulation Output

Validation is the process to determine whether the simulated model accurately replicates the real system or not. In this process, different conditions than the calibration process are simulated and the accuracy is checked with the real data. To be a valid model, the difference

of simulated and actual data should be below a predefined threshold. In this study, the threshold is taken as 10% variation of simulated speed from actual speed. If the model under study is valid, it can be safely assumed that it correctly replicates the real world situation even though it has been calibrated using a different dataset. Figure 5.8 shows the validation framework.

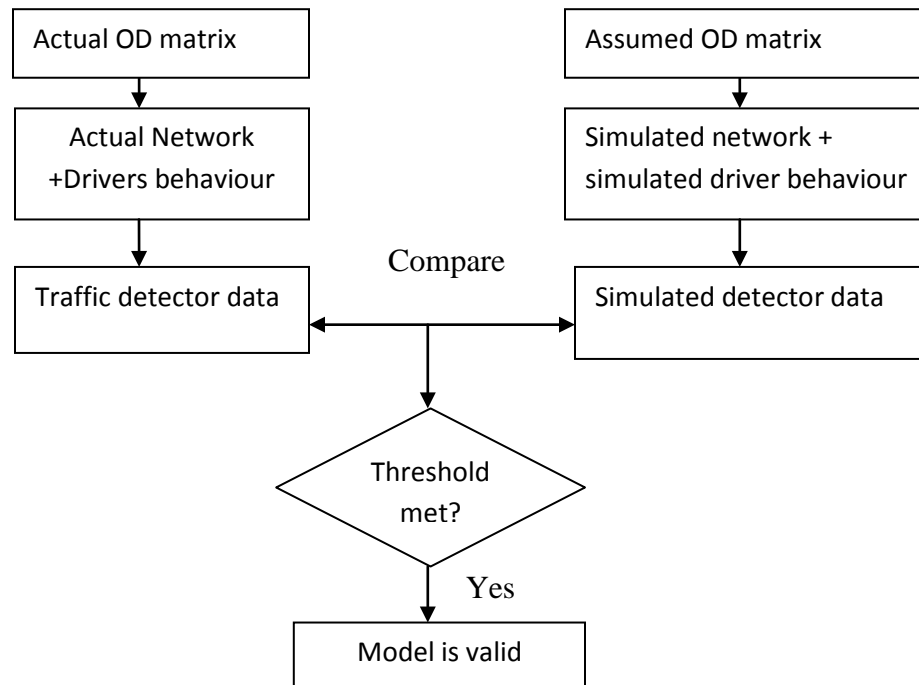


Figure 5.8: Framework for validation process

5.4.1 Selected Parameter for Validation

Similar to the calibration process, vehicle speed rather than traffic flow is selected as the parameter for validation.

For the validation of the model, we selected a time period when an accident has occurred on the road network. Although the simulation calibration has been done under normal traffic conditions but if it has been correctly calibrated, it should be able to replicate the traffic situation even during and after the incident as well. This is especially relevant as the main purpose of developing this simulation model is to use it for testing traffic mitigation schemes which can reduce the crash hazards.

5.4.2 Validation Results

For validation, following days and times are selected because an accident has occurred on these days during the simulation time in the corresponding direction:

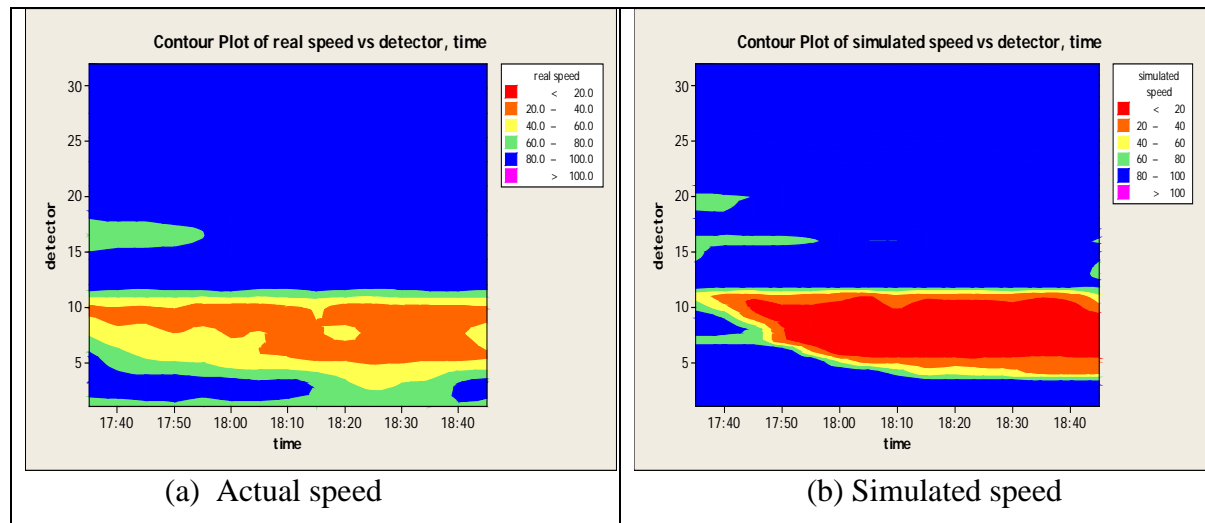
- Outbound direction: date: 16/11/2009, time: 17:35 to 18:45
- Inbound direction: date: 11/06/2009, time: 17:50 to 18:55

The simulated speed, actual speed and difference between these values for all the sections of the roadway are presented in contour plots. In the contour plot, speed ranges are kept as 20 km/hr in case of actual and simulated speed illustrations whereas ranges of 10 km/hr are used for the illustration of speed difference.

It is observed that the simulated speed values used for validation are very close to the actual speed values. Most of the values are within an error range of 10% while some values are nearly similar to the actual values. So, it can be said that the simulation model is well validated. The following figures show the results of the validation.

Case 1: Outbound direction (16/11/2009)

It is observed from the Figure 5.9 (c) that the difference in simulated speed and actual data is less than 10 km/hr for most of the sections. However, in few sections, the difference is in the range of 10 km/hr to 20 km/hr. So, the model can be considered as a valid model.



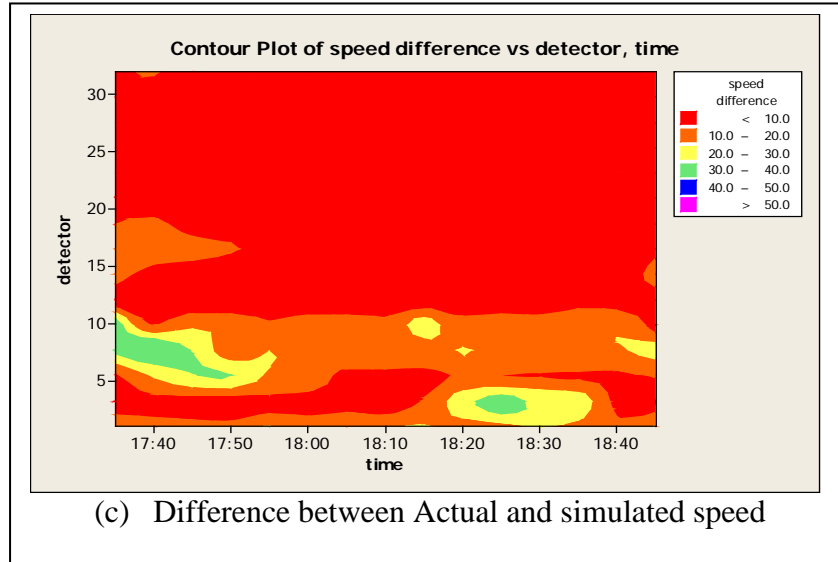
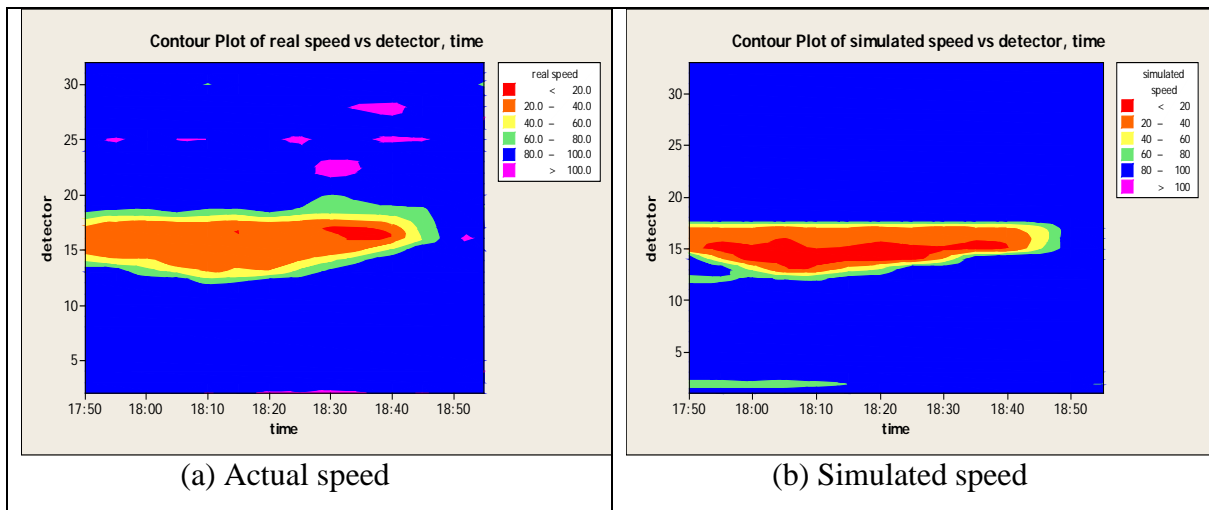


Figure 5.9: Validation result for outbound sections on 16/11/2009

Case 2: Inbound direction (11/06/2009)

In this case, the difference in simulated speed and actual data is less than 10 km/hr for most of the sections. However, in few sections, the difference is in the range of 10 km/hr to 20 km/hr. Thus, the model can be said as a validated model. The following figures represent the overall results.



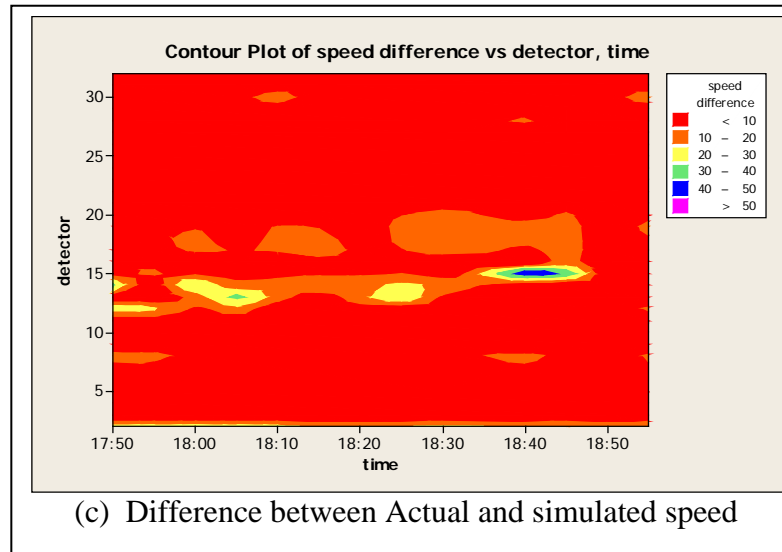


Figure 5.10: Validation result for inbound sections on 11/06/2009

5.5 Summary

The results from the Calibration and validation process show that the simulation model has the ability to replicate the real world traffic network. So the developed simulation model can be used to assess the detection and mitigation of crash hazards on the road.

Chapter 6

APPLICATION OF MITIGATION SCHEMES FOR HAZARDOUS CONDITION

Crash probability model is implemented with the aid of simulation model to detect the location of the hazardous condition in the roadway. Based on that detection, different mitigation schemes are applied to alleviate the adverse condition. This chapter describes the selected mitigation schemes and their application as well as effectiveness to reduce the hazardous condition on the roadway.

6.1 Reasons for Applying Mitigation Scheme

Road hazard has several impacts on traffic performance. It is one of the major causes of delays. Due to recurrent congestion, travel time on a road network become less predictable. It also causes lost productivity. Furthermore, hazardous conditions in roadways have environmental consequences also. Tailpipe emissions especially carbon monoxide and hydrocarbons emission are inversely related to speed. This means congested traffic causes more pollution than normal traffic condition (Deloitte, 2003).

As per Luk et al. (2008), the possible reasons for congestion on a freeway include:

- Traffic flow in the through sections of the freeway in excess of capacity.
- Uncontrolled access from on-ramps to through lane.
- Inadequate road geometries.
- Lack of real-time information.
- Frequent and unnecessary lane changing and speed changes.

The above mentioned causes of congestion can be eliminated by applying traffic control and management strategies. Thus, in order to alleviate the adverse effects of congestion on the freeway, mitigation schemes need to implement on freeway.

6.2 Different Mitigation Schemes

Several control and management strategies have been employed to improve roadway safety and to reduce hazardous condition. As this study focuses traffic condition of freeway, the control strategies that are applied in freeway are discussed here. The following control tools are used in order to provide safety and efficiency of freeway traffic (Luk et al., 2008):

- Automatic Incident Detection (AID) system
- Ramp metering
- Variable Speed Limit (VSL)
- Driver Information System (DIS)
- Traffic lane management system

Automatic incident detection system is employed to detect the non-recurrent congestion or accidents after they have occurred. Then traffic management centre applies the necessary management system (Luk et al., 2008).

Ramp metering strategy is used on the freeway to reduce congestion. In this system, traffic signals are set at the freeway on-ramps in order to control the entry of the vehicle from on-ramps to freeway sections (Scariza et al., 2003).

Variable speed limit is such a control strategy that can evaluate the appropriate speed of vehicle to travel through the roadway. This evaluation is done based on traffic condition, weather information and road surface condition. The regulatory speed is displayed through variable message sign (Jiang et al., 2011)

Driver information system is applied through variable message signs (VMS), radio broadcasting, internet access, highway radio or personal services. The information includes changing travel mode before starting a trip, changing lane, reducing speed or changing destinations etc.

Traffic lane management system is used in congested site during peak hours. The main objective of this system is to increase the capacity of the roadway. It includes contra flow operation, using shoulder lane for traffic movements etc.

6.3 Selected mitigation schemes

The crash risk probability model developed in this study is useful to prevent roadway congestion propagation. So, the mitigation scheme should have such capability to prevent congestion before they occur. Ramp metering and variable speed limit strategy have the ability to do this. However, AID system is effective after incidents have occurred (Luk et al., 2008). Except VMS, other features of the driver information system are not applied in this study since variable message signs (VMS) are expected to have better performance than radio broadcasting, internet access, highway radio etc. If there were no such limitation, this study would not. The VMS system used here is one kind of variable speed limit strategy (Luk et al., 2008). And traffic lane management system is not applicable for the selected freeway.

Thus, among the mitigation schemes, two are selected for this study in order to control traffic performance to mitigate hazardous condition in the freeway. These are:

- Ramp Metering strategy
- Variable Speed Limit scheme

6.3.1 Ramp Metering

Ramp metering is a special traffic control system installed on the on-ramps of freeway or certain roads to regulate traffic flow (TSS-Transport Simulation Systems, 1997). This is a strategy of using the traffic signals at freeway on-ramps to manage the rate of vehicles entering the freeway. This system is designed to improve the average speed of the vehicles on the freeway through lanes by alleviating the congestion developed by either road incidents or excessive traffic flow (State of California Department of Transportation, 2005). Although original ramp meters were pre-timed, now-a-days, nearly all the ramp meters are sensitive to real time traffic condition (Scariza et al., 2003).

Ramp metering scheme has following positive outcomes (Scariza et al., 2003):

- Improves the condition of flow on freeway mainline.
- Decreases the rate of incidents, fuel consumption and emission.
- Breaks up platoons to allow more efficient merging.

- Alters network routings to gain greater balance and efficiency.

Ramp meter has disadvantages also (Scariza et al., 2003):

- Causes huge congestion and delays on the ramps in many cases.
- Causes network re-routing which has negative impact on alternate routes.

6.3.2 Variable Speed Limit (VSL)

Variable speed limit scheme is a type of traffic control strategy which applies the regulatory speed based on roadway and traffic condition in a dynamic manner. The main objective of VSL is to maintain safety or to increase capacity on roadways (Bevrani et al., 2011). This control system is employed in real time situations such as traffic condition, adverse weather condition, road surface condition and work zones (Warren, 2003). VSL strategy typically includes sensor, variable speed limit sign, variable message sign (VMS) and a central processing unit. Depending on the roadway traffic condition, the central processing unit computes time varying appropriate speed limit and display it through VMS (Lin et al., 2004; Jiang et al., 2011).

Variable speed limit has the following potential advantages (Luk et al., 2008):

- Reduces flow breakdowns
- Reduces road crashes
- Increases freeway efficiency
- Provides general speed management during planned events such as road works or sports events.

Variable speed limit has disadvantages also:

- Increases travel time.

6.3.3 Effectiveness of the Selected Mitigation Strategies

The effectiveness of ramp metering and variable speed limit strategy can be assessed in context of mitigation congestion and improving traffic safety. Section 2.4 in chapter 2 discusses the effectiveness of these two mitigation scheme in details.

6.4 Application of Mitigation Scheme in Simulation Software (AIMSUN)

AIMSUN 6.0 provides different types of ramp metering schemes and variable speed limit strategy.

6.4.1 Application of Ramp Metering in AIMSUN

6.4.1.1 Types of Ramp Metering

AIMSUN 6.0 can implement three types of ramp metering strategy (TSS-Transport Simulation Systems, 1997). These are:

Green time metering: The ramp metering is controlled by a traffic signal that turns red and green on a cyclical basis. If we are using fixed traffic control here, there will be unique green time. If the metering control is external, the minimum and maximum green fields show the minimum and maximum values for the acceptable range of green time variation. For the rest of the cycle time, the traffic signal will be red. Yellow time is also modelled in green time metering. (TSS-Transport Simulation Systems, 1997)

Flow metering: In this case, the ramp metering objective is to let a certain number of vehicles per hour go through the meter. Each time the meter is opened to release vehicles, it is done in such a way that platoons, whose length is "platoon length" parameters, can pass. If the control of the metering is fixed, there will be only one parameter to specify the number of vehicles per hour to be released. If the control is external, also the minimum and maximum flow determining the acceptable range of flow variation will be specified. (TSS-Transport Simulation Systems, 1997)

Delay metering: This type of metering is used to model tolls, customs, checkpoints or any other type of individual control. It is assumed that every vehicle will have to stop at the control point for a certain time. This time is assumed to be a normally distributed random

variable with a mean delay time and standard deviation (TSS-Transport Simulation Systems, 1997).

6.4.1.2 Types of the Control

There are four types of control policies associated to the ramp metering (TSS-Transport Simulation Systems, 1997) which are as follows:

Unspecified: When no control editing has been made.

Uncontrolled: When the metering device is assumed to be disconnected.

Fixed: When a fixed criterion for allowing flow to go through is used which means none of the parameters of the metering will vary in time for this control.

External: The parameters initially set can vary between a given maximum and minimum depending on traffic conditions and according to an external user-set algorithm (TSS-Transport Simulation Systems, 1997).

6.4.1.3 Parameter Selection

Depending on the ramp metering type, different parameters need to be defined in AIMSUN.

For green time metering, the user must set the cycle's total time, that is the time of green plus the time of red, the yellow time, the offset of the first green time related to the initial time of the control plan, the time of green and if the control type is external a minimum and a maximum time of green (TSS-Transport Simulation Systems, 1997).

For flow metering, the user must set the desired flow in vehicles per hour that is to be achieved with the metering. The platoon of vehicles specified for the metering will indicate the maximum number of vehicles that will go through the metering during each green (TSS-Transport Simulation Systems, 1997).

For delay metering, the user must set the metering mean delay for each vehicle and the standard deviation from the specified mean (TSS-Transport Simulation Systems, 1997).

6.4.2 Application of Variable Speed Limit System in AIMSUN

Variable speed limit is applied in the through section of the roadway. In AIMSUN, although a speed is applicable for each section, it can be changed by using ‘speed change’ option. There additional information like vehicle class and changed speed should be mentioned there. Vehicle class may be fixed as all type of vehicle or any certain type of vehicle, and speed can be changed from 0 to higher than 100 km/hr.

6.5 Traffic Scenarios Analysis Method

In this study, traffic scenarios during no-mitigation and during mitigation are assessed. For analysis of the scenarios, four real world situations have been replicated in simulation procedure. The procedure of the scenario analysis is as follows:

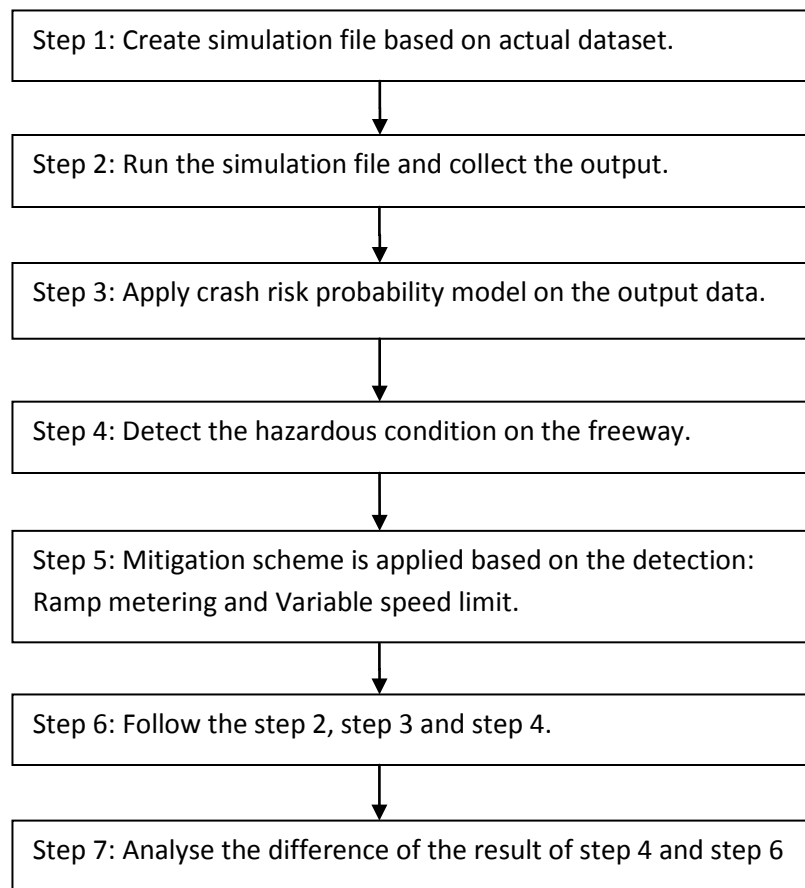


Figure 6.1: Scenario analysis steps

1. First, the simulation file is created based on real dataset by considering all the factors. The total duration of the simulation file depends on actual accident occurrence time. The duration is either five hours or six hours: two hours before the accident situation and three or four hours after accident.
2. After running the simulation file, traffic parameters data i.e. traffic flow (Flow SRC) and average vehicle speed (Speed SRC) data for every sections in the particular direction of the freeway are collected from the output table.
3. To find out hazardous conditions from the simulation output, crash risk probability model is applied to this dataset. This gives the indication of the probable hazardous and non-hazardous situation. The indication consists of two values 1 and 0. The value 1 is shown for hazardous condition and the value 0 is shown for non- hazardous condition.
4. Then mitigation scheme is applied based on the first hazardous condition detected by the model. Mitigation system initialised at five minutes after the first detected hazardous condition. Two mitigation schemes: ramp metering and variable speed limit are applied separately. Ramp metering scheme is applied at the upstream on-ramps of the detected location and variable speed limit strategy is applied on the upstream through lanes. Then outcomes from each strategy are analysed.
5. Next, crash risk probability model is again applied to the output from the simulation file. This is done for both mitigation schemes. The outcomes from this step are compared with that of no- mitigation condition (the result of step 3).
6. Finally, It is checked whether the mitigation schemes improve the traffic condition or not. Travel time or speed data can be the improvement measurement criterion. For this study, vehicle speed is considered as the key criterion for measuring anticipated improvement because of data availability.

6.6 Application of Ramp Metering Scheme

In this study, ramp metering is adopted in order to mitigate hazardous conditions in the freeway. The following ramp metering criteria has been selected in AIMSUN:

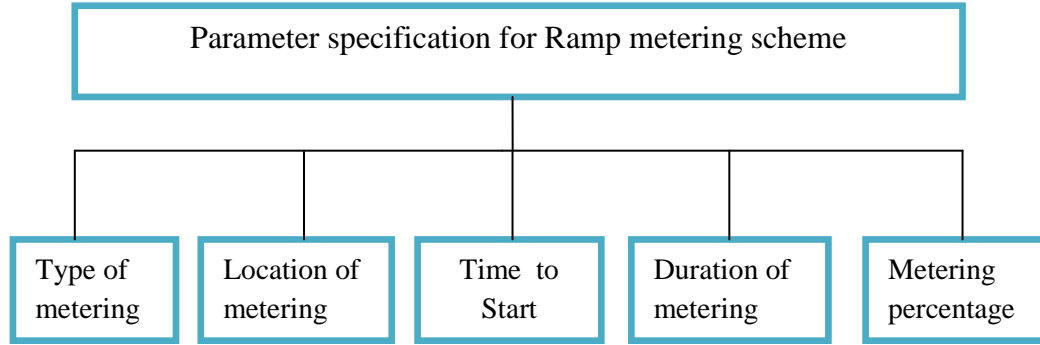


Figure 6.2: Selection of parameters for Ramp Metering strategy

Type of ramp metering: In this study, flow metering strategy is selected because flow rates of the on-ramps are included in the real dataset. Based on this flow rate, the desired flow for mitigation process can be implemented. Albeit, traffic flow can be metered in a good manner in the green metering process, the exact amount of flow rate reduction cannot be found. On the other hand, delay metering process is exclusively for tolls, customs and check point. So, for the convenience of the implementation, flow metering process is applied as the hazardous condition mitigation scheme.

Location of ramp metering: Ramp metering strategy is applied at the on-ramps of the freeway. Generally, traffic flow at the downstream on-ramps has no effect to improve the traffic condition of the upstream sections. Thus, in this study, ramp metering scheme is implemented at the upstream on-ramps of the hazardous location.

Time to install: For simulation purpose, the time of initialisation of ramp metering at on-ramps is set to five minutes after the detection of hazardous condition. This is not realistic if the ramp metering starts just after the detection of the hazard as it takes few minutes to detect and install the metering scheme.

Duration of the metering: The duration of the metering is selected as one hour after the initialisation. This time is obtained by analysing the real dataset. Real dataset includes duration of accident removal time and the time needed to regain the normal traffic flow after accident. It is observed in most cases, it takes about 50 minutes to revive normal traffic flow after hazardous situation. Additional 15 minute (as installed after 5 minutes of hazard) is considered in this study due to concern of special cases.

Metering percentage: One of the important issues of ramp metering scheme is to determine the percentage or the number of vehicles that needs to be metered. By applying different percentage of metering in different situations, the best percentage of ramp metering scheme can be found. This percentage needs to apply in order to get best traffic condition.

6.7 Application of Variable Speed Limit (VSL) Strategy

In this study, Variable Speed Limit is the second mitigation scheme explored for alleviation of hazardous conditions on the freeways. The following parameters need to be specified for VSL in AIMSUN 6.0:

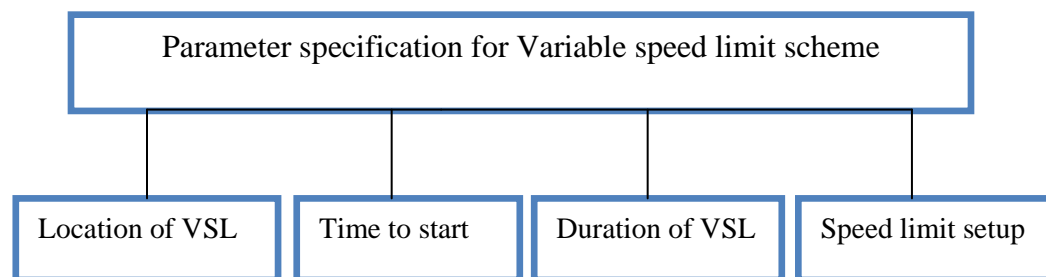


Figure 6.3: Selection of parameters for Variable Speed Limit strategy

Location of VSL: Variable speed limit scheme is applied for the through lanes/sections of the freeway. Generally, the downstream sections of the hazardous location have no effect to improve the traffic condition on the hazardous location. Thus, in the process of evaluating the mitigation scheme, VSL is installed at the upstream road sections of the hazardous location. One of the important issues in this mitigation strategy is to select the number of upstream sections where VSL needs to be applied. In this study, up to 1km, 2km and 3km upstream

sections of the hazardous location are considered to apply VSL scheme. By evaluating the mitigation results, the best one among these distances will be selected.

Time to install: In real situations, it takes few minutes to detect and install the VSL scheme. Thus, the time of initialisation of VSL at the upstream sections is set to five minutes after the detection of hazardous conditions.

Duration of VSL: The duration of VSL is selected as one hour after the initialisation. This time is obtained by analysing the real dataset. Real dataset includes duration of accident removal time and the time needed to regain the normal traffic flow after accident. It is observed in most cases, it takes about 50 minutes to revive normal traffic flow after hazardous situation. Additional 15 minute (as installed after 5 minutes of hazard) is considered in this study due to concern of special cases.

Set up speed limit: One of the important factors for VSL application is to select the regulatory speed for the duration of its application. By applying different speed in different situations, the best regulatory speed for VMS strategy can be found. This speed needs to be applied on the upstream of the detected hazardous condition section in order to minimise the impact.

6.8 Parameter used for ramp metering and VSL strategy

Two basic types of mitigation schemes, ramp metering and variable speed limit are applied in this study. It needs to be mentioned that both ramp metering and variable speed limit schemes are applied on the upstream sections of the hazardous location. This is due to the fact that controlling downstream traffic cannot improve hazardous conditions. Various percentages of ramp flow metering with different number of metered on-ramps and different speed limits on different upstream locations in VSL systems are employed to find out the suitable mitigation scheme for each of the mentioned traffic conditions.

In ramp metering process, three different percentages (90%, 70% and 50%) of metered traffic at three different number of nearest on-ramps (1, 2 and 3) are tested. For the VSL schemes, four speed limits . 90 km/hr, 80 km/hr, 70 km/hr and 60 km/hr are applied on three different upstream section lengths: 1 km, 2km and 3 km from the hazard location.

In this study, 50% ramp metering is selected as the lowest percentage of metering because it is observed from the simulation trials that on-ramps become congested if the metering percentages are lesser than this. Similarly, 60 km/hr is taken as the lowest speed limit for VSL system by considering the travel times through the sections.

6.9 Effects of mitigation schemes to reduce hazardous situations

For this study, the efficiencies of the mitigation schemes are evaluated by the reduction percentages of hazardous conditions that developed during no-control situation. For assessing the mitigation schemes, three different traffic conditions are analysed in the simulation process. These are:

- Peak hour traffic condition
- Low volume traffic condition
- High volume traffic condition

Traffic condition during afternoon peak hour between 4 p.m. to 8 p.m. is considered as the peak hour traffic condition since the mitigation schemes are tested on the outbound direction. The low and high traffic conditions are taken as 50 % of peak hour traffic and 125% of peak hour traffic respectively.

Case 1: For peak hour traffic condition

The effects of mitigation schemes in peak hour traffic are analysed using micro-simulation process. A summary of the simulation process features is mentioned below:

- Traffic direction: *Outbound*
- Date: *16/11/2009* and Time: *4 p.m. to 9 p.m.*
- First hazardous condition detected by the model: *4:10 p.m. on section 16*
- Available upstream on-ramp: *2,7*
- Available upstream sections: *1 to 15*
- For ramp metering scheme:
 - Initialisation time: *4:15 pm*
 - Duration: *1 hour*
 - Metering percentage: *50%, 70% and 90%*

- Applied at on-ramp sections:
 - Nearest one on-ramp:7
 - Nearest two on-ramps:2, 7
- For variable speed limit scheme:
 - Initialisation time: 4:15 p.m.
 - Duration: 1 hour
 - Speed limits: 90 km/hr, 80 km/hr, 70 km/hr and 60 km/hr
 - Applied on sections:
 - 1 km upstream: section no.16,15,14
 - 2 km upstream: section no.16,15,14,13,12
 - 3 km upstream: section no.16,15,14,13,12,11,10

■ ***During no-control situation:***

Once the traffic simulation has been run under no-control situation, the crash risk probability model is applied to evaluate the crash hazard conditions. Figure 6.4 shows the hazard condition of the freeway in no-control situation. Here, hazard=1 represents hazardous condition and hazard=0 represents no-hazard condition. It is observed that during no-control condition, the model has found total 79 probable hazards during the simulated time period on this roadway.

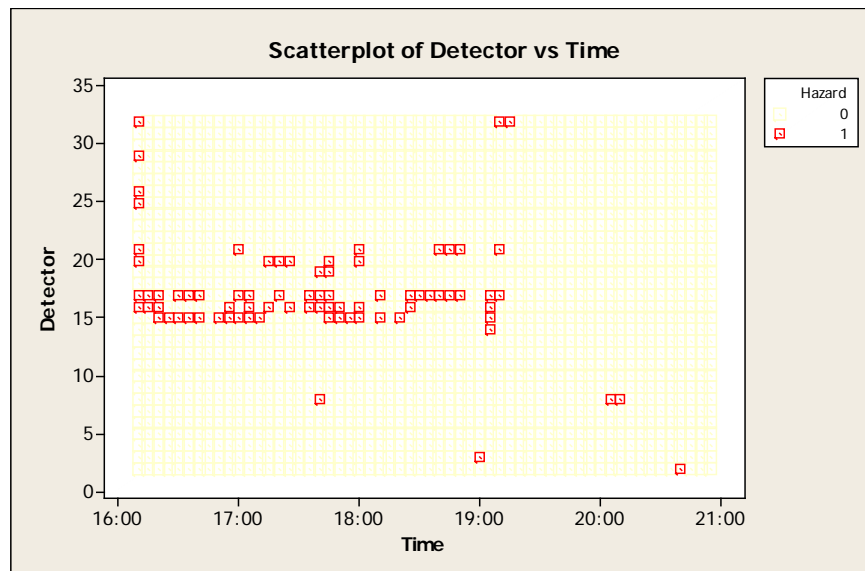


Figure 6.4: Time-space diagram during no-control situation in peak hour traffic

▪ ***After implementing Ramp Metering scheme:***

After applying different types of ramp metering strategies, a significant number of hazardous conditions that appeared during no-control situation have been removed. Table 6.1 shows the overall results of the probable hazards under different metering percentages and number of metered ramps.

Table 6.1: Improvements of hazard condition after ramp metering in peak hour traffic

Metered percentage \ No. of ramps metered	50% metering		70 % metering		90% metering	
	Total hazard	Hazard reduction	Total hazard	Hazard reduction	Total hazard	Hazard reduction
1	72	8.86%	67	15.18%	72	8.86%
2	58	26.58%	74	6.32%	72	8.86%

It is observed from the above table that the highest improvements in traffic condition can be found after applying 50% ramp metering scheme at the nearest two upstream on-ramps. However, the result shows that only one nearest on-ramp with 70% ramp flow metering provides more improvements than two on-ramps metering. This is due to the fact that this percentage of ramp metering influences the traffic condition to become more prone to hazard. But the same result comes out after several simulations. It is seen that in 50% ramp metering scheme, increase in the number of metered ramps results in further improvement. In case of 90% ramp metering, the hazard reduction percentages are same for both one and two on-ramp metering. This is because of the less percentage of metering which may not have significant impacts on through traffic reduction and hence the hazard reduction. Although all the mentioned ramp metering strategies show reduction in hazardous conditions, no specific trend can be drawn. Figure 6.5 illustrates the overall improvements in hazardous conditions by different ramp metering schemes:

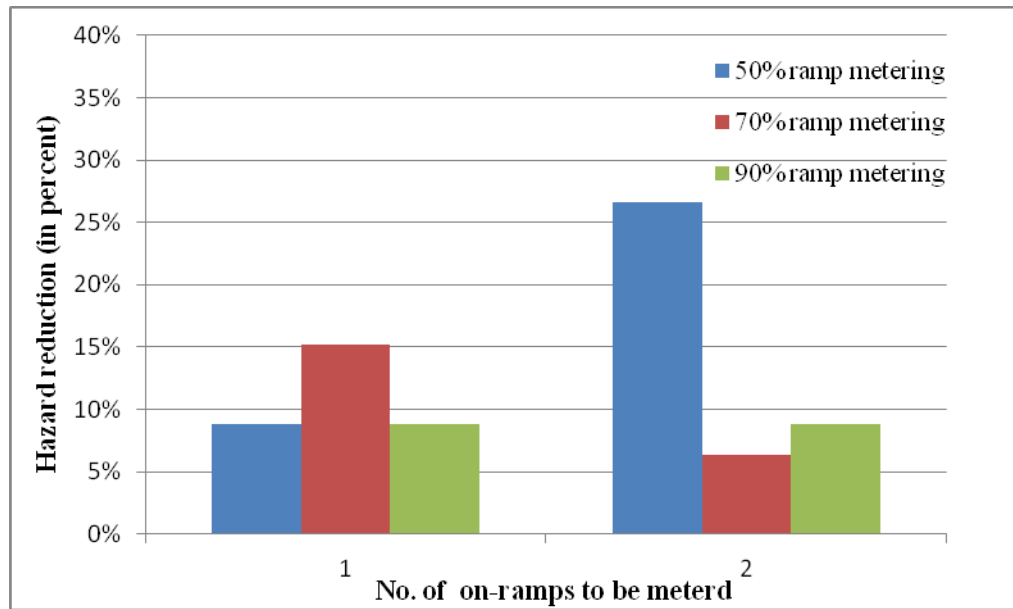


Figure 6.5: Improvements of hazards by ramp metering schemes in peak hour traffic

Figure 6.6 shows the hazard condition of the freeway after applying 50% ramp metering at two nearest upstream location of hazard. It is observed that after applying this ramp metering scheme, the model has indicated total 58 probable hazards during the simulated time period which is much less than that of no-control situation.

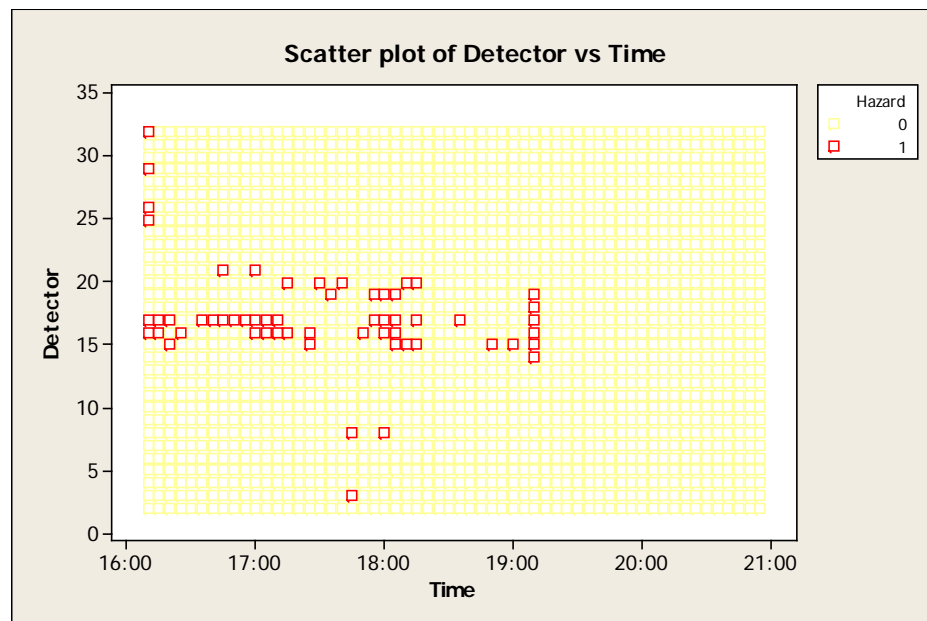


Figure 6.6: Time-space diagram of hazardous condition by ramp metering in peak hour traffic

▪ ***After implementing VSL scheme:***

Implementation of variable speed limit strategies also provides less hazardous conditions in peak hour traffic. The following table shows the improvements made by variable speed limit schemes.

Table 6.2: Improvements of hazard condition after VSL schemes in peak hour traffic

Speed limit \ Distance from location	60 km/hr		70 km/hr		80 km/hr		90 km/hr	
	Total hazard	Hazard reduction	Total hazard	Hazard reduction	Total hazard	Hazard reduction	Total hazard	Hazard reduction
1 km	52	34.17%	64	18.98%	58	26.58%	66	16.45%
2 km	65	17.72%	54	31.64%	57	27.84%	54	31.64%
3 km	75	5.05 %	57	27.84%	68	13.92%	71	10.12%

Table 6.2 shows that the variable speed limit strategy can reduce hazardous conditions more effectively if it is applied within one or two kilometre upstream of the hazard location. The highest improvement can be found by applying the speed limit of 60 km/hr. However, variable speed limit installed on a 3 kilometres long upstream section of the hazard location does not improve it further. Figure 6.7 represents the reduction percentages of hazard after applying each of the VSL strategies.

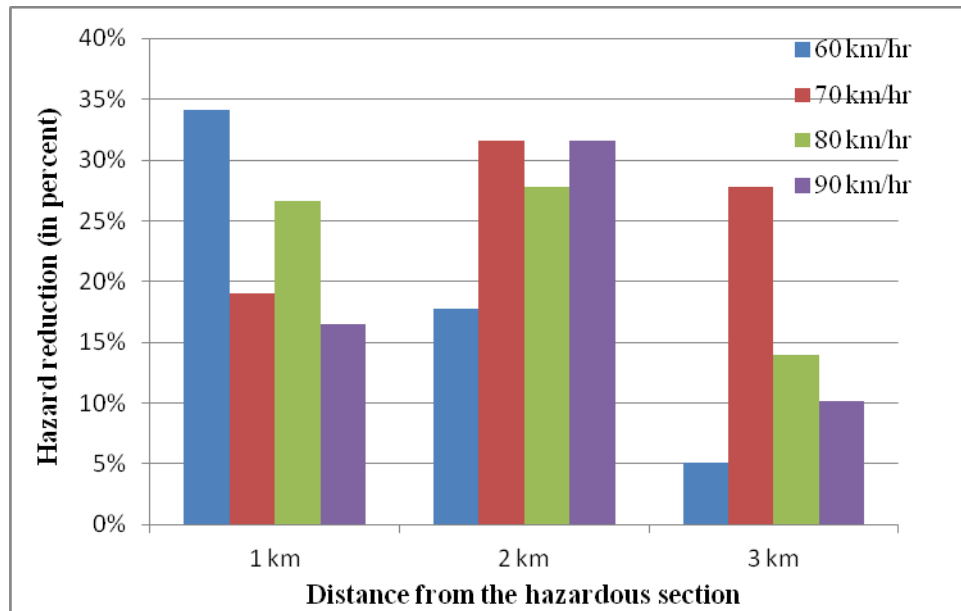


Figure 6.7: Improvements of hazards after VSL schemes in peak hour traffic

Figure 6.8 shows the hazard condition of the freeway during variable speed limit scheme. It is observed that, the model has indicated total 52 probable hazards during the simulated time period after applying VSL at 1 km upstream section of hazard. It is observed that after applying this strategy, the number of hazards reduces significantly.

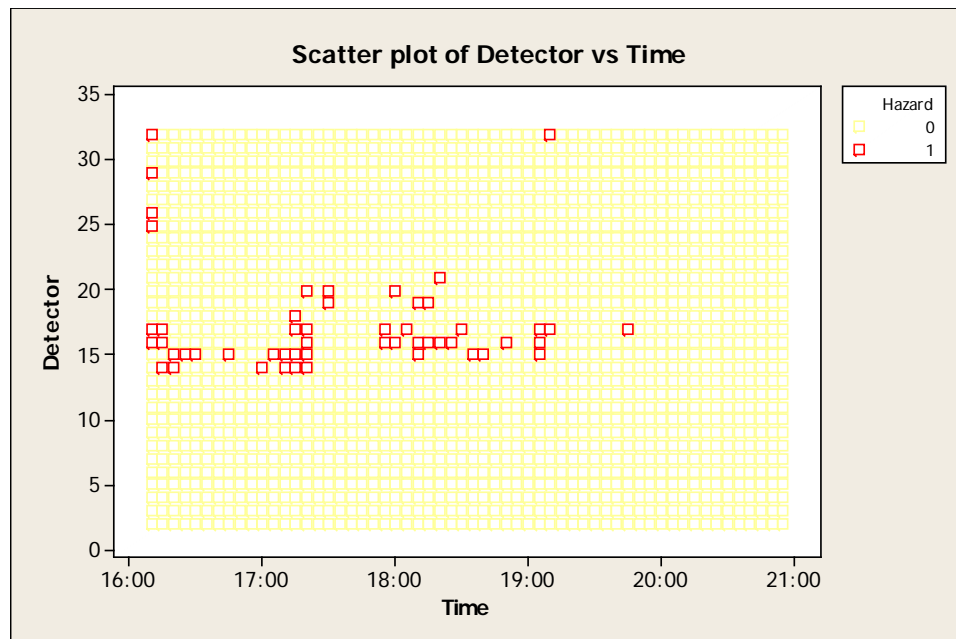


Figure 6.8: Time-space diagram of hazardous condition by VSL in peak hour traffic

Case 2: For low volume traffic condition

The impacts of mitigation schemes in low volume traffic condition are also assessed using the traffic simulation on the same stretch of the road albeit with 50% of the peak hour traffic demand. The features of the simulation process are as follows:

- Date: *16/11/2009* and Time: *4 pm to 9 pm*
- Traffic direction: *Outbound*
- First hazardous condition detected by the model: *4:10 pm on section 26*
- Available upstream on-ramp: *2, 7, 16, 20, 24*
- Available upstream sections: *1 to 25*
- For ramp metering scheme:
 - Initialisation time: *4:15 p.m.*
 - Duration: *1 hour*
 - Metering percentage: *50%, 70% and 90%*
 - Applied at on-ramp sections:
 - Nearest one on-ramp: *24*
 - Nearest two on-ramps: *20, 24*
 - Nearest three on-ramps: *16, 20, 24*
- For variable speed limit scheme:
 - Initialisation time: *4:15 p.m.*
 - Duration: *1 hour*
 - Speed limits: *90 km/hr, 80 km/hr, 70 km/hr and 60 km/hr*
 - Applied on sections:
 - 1 km upstream: section no. *26, 25, 24*
 - 2 km upstream: section no. *26, 25, 24, 23, 22*
 - 3 km upstream: section no. *26, 25, 24, 23, 22, 21*
- ***During no-control situation:***

Figure 6.9 illustrates the hazardous and non-hazardous condition of the freeway in no-control situation in which hazard=1 represents hazardous condition and hazard=0

represents no-hazard condition. It is observed that during no-control situation, the model has indicated 20 probable hazards during the above mentioned time in this roadway.

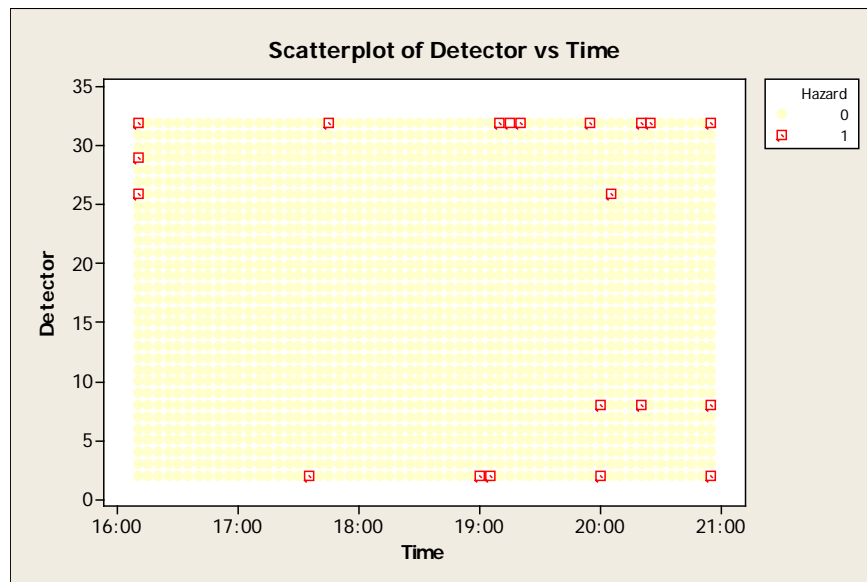


Figure 6.9: Time-space diagram during no-control situation in low volume traffic

▪ ***After applying Ramp Metering scheme:***

Application of ramp metering process during low volume traffic shows reduction in hazardous conditions. Table 6.3 describes the improvements made by different ramp metering systems.

Table 6.3: Improvements in hazards by ramp metering schemes in low volume traffic

Metered percentage \ No. of ramp metered	50% metering		70 % metering		90% metering	
	Total hazard	Hazard reduction	Total hazard	Hazard reduction	Total hazard	Hazard reduction
1	16	20.0%	16	20.0%	16	20.0%
2	15	25.0%	13	35.0%	15	25.0%
3	12	40.0%	13	35.0%	15	25.0%

This table depicts that the highest reduction percentage of hazardous condition is found after applying 50% ramp metering at nearest three upstream on-ramps of the hazard location. It is noteworthy that in 70% and 90% ramp metering, the 2 and 3 on-ramps results are similar

while for 50% ramp metering, 3 on-ramps condition provides 15% more reduction of hazards than for 2 on-ramps situation. So it is observed that in all percentage of ramp metering schemes, the higher improvements in removing hazard are found when all the nearest three on-ramps are metered. However, the hazard reduction percentages are same for all types of on-ramp metering strategy. The Figure 6.10 represents the outcomes of the application of variable speed limit strategies.

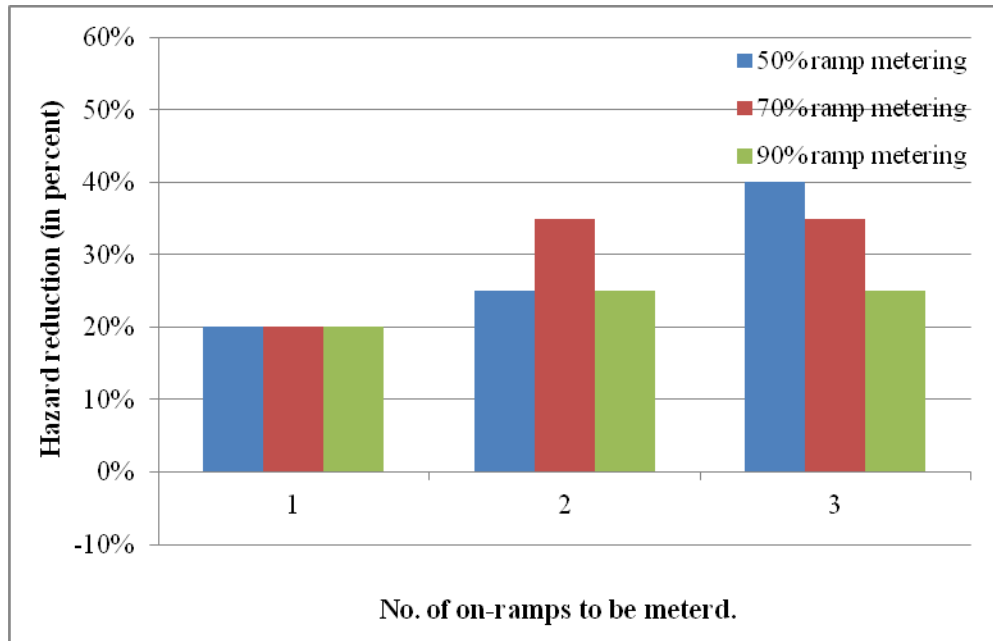


Figure 6.10: Improvements of hazards by ramp metering schemes in low volume traffic

Figure 6.11 shows the hazard condition of the freeway after applying 50% ramp metering at three nearest upstream ramps of the hazard. It is observed that after applying this ramp metering scheme, the model has indicated total 12 probable hazards during the simulated time period which means hazardous conditions are reduced significantly by this scheme.

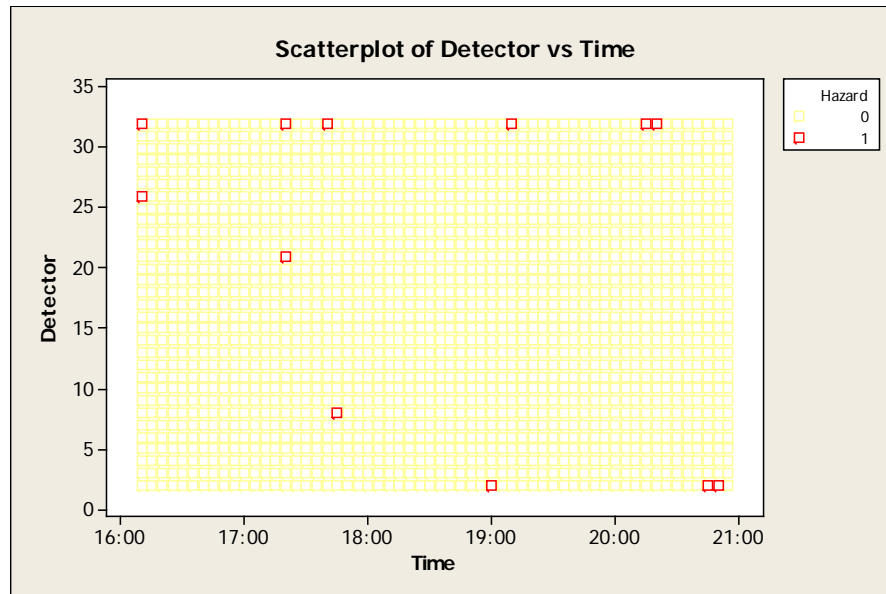


Figure 6.11: Time-space diagram of hazardous condition by ramp metering in low volume traffic

▪ *After applying VSL scheme:*

Similar to the ramp metering, implementation of variable speed limits also provides the improved traffic than that of no-control situation. The following table shows the overall results after implementing the VSL scheme.

Table 6.4: Improvements in hazard condition after VSL schemes in low volume traffic

Speed limit \ Distance from location	60 km/hr		70 km/hr		80 km/hr		90 km/hr	
	Total hazard	Hazard reduction	Total hazard	Hazard reduction	Total hazard	Hazard reduction	Total hazard	Hazard reduction
1 km	18	10.0%	14	30.0%	14	30.0%	11	45.0%
2 km	12	40.0%	17	15.0%	16	20.0%	15	25.0%
3 km	11	45.0%	12	40.0%	15	25.0%	15	25.0%

The table shows that the highest reduction of road hazard can be found by using variable speed limit of 60 km/hr up to 3 km long upstream sections of the hazard location. However, if the same speed limits are applied up to 1 km upstream sections, the reduction percentages are found to be very low. But the other speed limits such as 70 km/hr, 80 km/hr and 90 km/hr provide better results if installed within 1km upstream instead of 2 or 3 km upstream sections. The following chart illustrates the overall results.

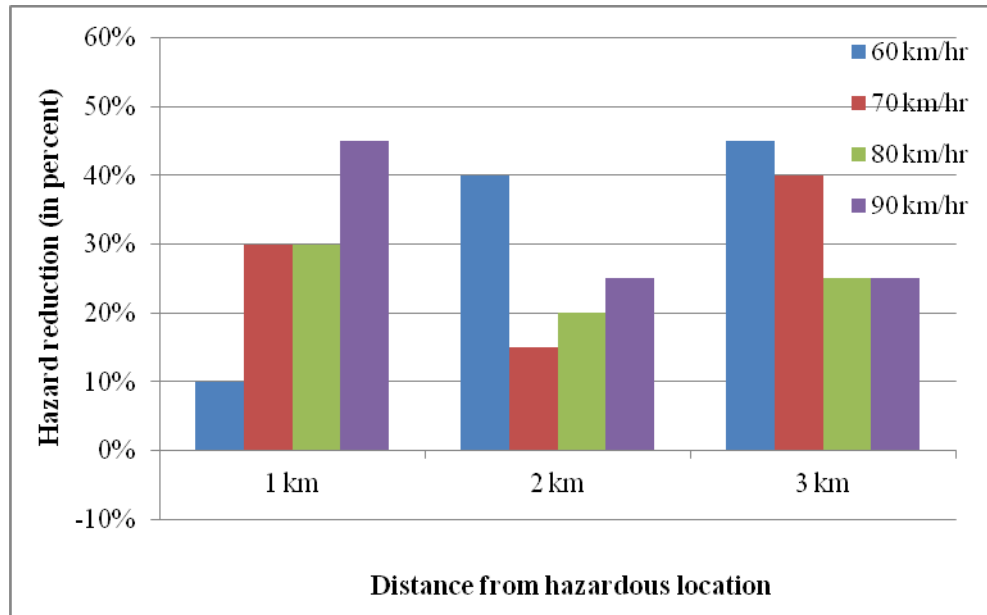


Figure 6.12: Improvements of hazards after VSL schemes in low volume traffic

Figure 6.13 shows the hazard condition of the freeway during variable speed limit scheme. It is observed that, the model has indicated 11 probable hazards during the simulated time period after applying VSL at 3 km upstream section of the hazard. It is seen from the figure that the first few hazards have shifted to next time interval and nearer location. The number of hazards is less than that of no-control situation.

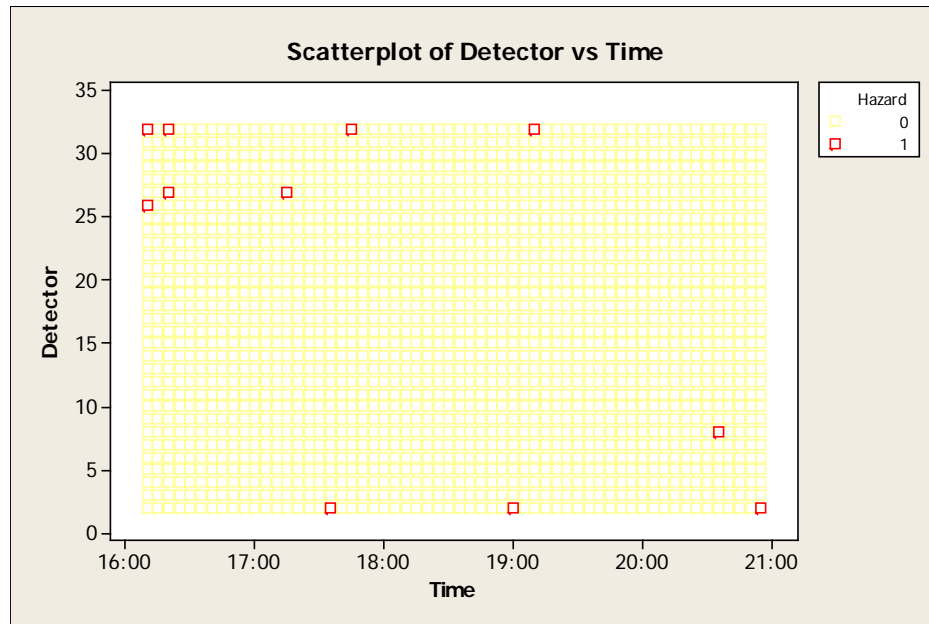


Figure 6.13: Time-space diagram of hazardous condition by VSL in low volume traffic

Case 3: For high volume traffic condition

The effects of mitigation schemes in high volume traffic condition are analysed using micro-simulation process. The overview of the features of the simulation process is described below:

- Date: *16/11/2009* and Time: *4 pm to 9 pm*
- Traffic direction: *Outbound*
- First hazardous condition detected by the model: *4:10 pm on section 8*
- Available upstream on-ramp: *2, 7*
- Available upstream sections: *1 to 7*
- For ramp metering scheme:
 - Initialisation time: *4:15 p.m.*
 - Duration: *1 hour*
 - Metering percentage: *50%, 70% and 90%*
 - Applied at on-ramp sections:
 - Nearest one on-ramp: *7*
 - Nearest two on-ramps: *2, 7*

- For variable speed limit scheme:
 - Initialisation time: 4:15 pm
 - Duration: 1 hour
 - Speed limits: 90 km/hr, 80 km/hr, 70 km/hr and 60 km/hr
 - Applied on sections:
 - 1 km upstream: section no. 8, 7, 6
 - 2 km upstream: section no. 8, 7, 6, 5, 4
 - 3 km upstream: section no. 8, 7, 6, 5, 4, 3

In this study, the mitigation scheme is applied based on the first detected hazardous location. If multiple hazardous locations are found at the same time, then applying mitigation strategies would be similar to the single case. This is caused by the limitation in the testing methodology used in this study. The methodology is not be applied at multiple locations at the same time since it is not based on online i.e. applying, changing, testing, re-applying the mitigation scheme and so on. Rather it is based on the procedure such as run the simulation then get the outputs and after that, make changes to show the mitigation and re-run the simulation.

▪ ***During no-control situation:***

Figure 6.14 shows the hazardous and non-hazardous condition of the freeway in no-control situation where hazard=1 represents hazardous condition and hazard=0 represents no-hazard condition. It is observed that during no-control situation, the model has found total 158 probable hazards during the above mentioned time in this roadway.

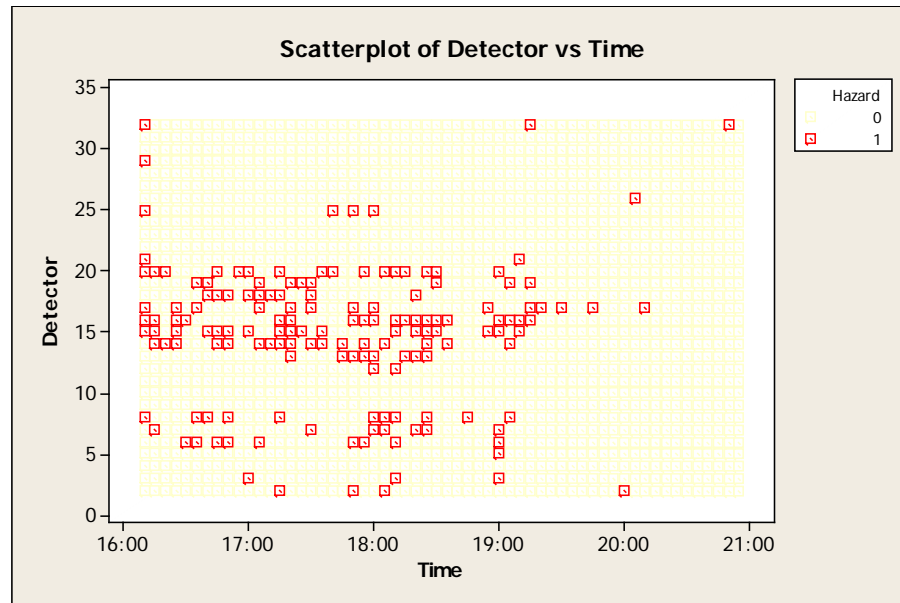


Figure 6.14: Time-space diagram during no-control situation in high volume traffic

▪ ***After implementing Ramp Metering scheme:***

After implementing different ramp metering strategies during high volume traffic, a significant number of hazardous conditions that appeared during no-control situation have been removed.

Table 6.5: Improvements of hazard condition by ramp metering in high volume traffic

Metered percentage \ No. of ramp metered	50% metering		70 % metering		90% metering	
	Total hazard	Hazard reduction	Total hazard	Hazard reduction	Total hazard	Hazard reduction
1	126	20.3%	126	20.3%	130	17.7%
2	126	20.3%	129	18.4%	135	14.5%

It can be concluded from the above results that in heavy traffic conditions, the hazard reduction percentages by various ramp metering strategies are quite similar regardless of the number of upstream metered on-ramps. However, in 90% ramp metering strategy, the percentage of hazard reduction is slightly lower than that of the other two. This may happen

because of less reduction percentage of on-ramp's traffic volume. Compared to two metered ramp scheme, one metered ramp strategy provides better traffic condition. This is due to the fact that additional percentage of ramp metering has effect on traffic parameters by influencing the traffic condition to become more prone to hazard. But same output has been found from different simulation runs. The following figure shows the overall hazard reduction percentages achieved by each ramp metering process.

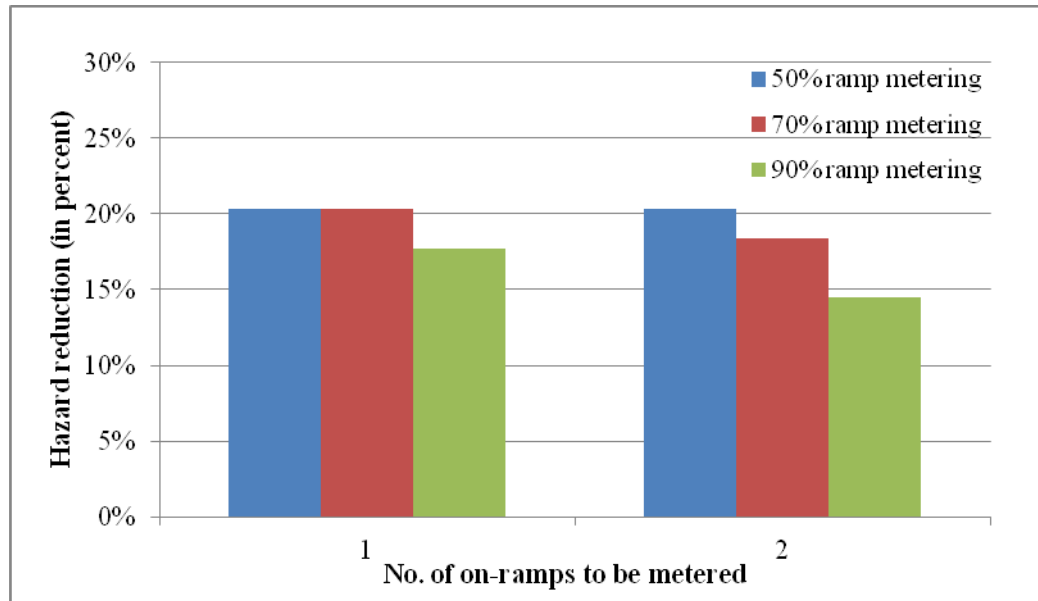


Figure 6.15: Improvements of hazards by ramp metering schemes in high volume traffic

Figure 6.16 shows the hazard condition of the freeway after applying 50% ramp metering at two nearest upstream location of the hazard. It is observed that after applying this ramp metering scheme, the model has indicated total 126 probable hazards during the simulated time period which means number of hazards have reduced.

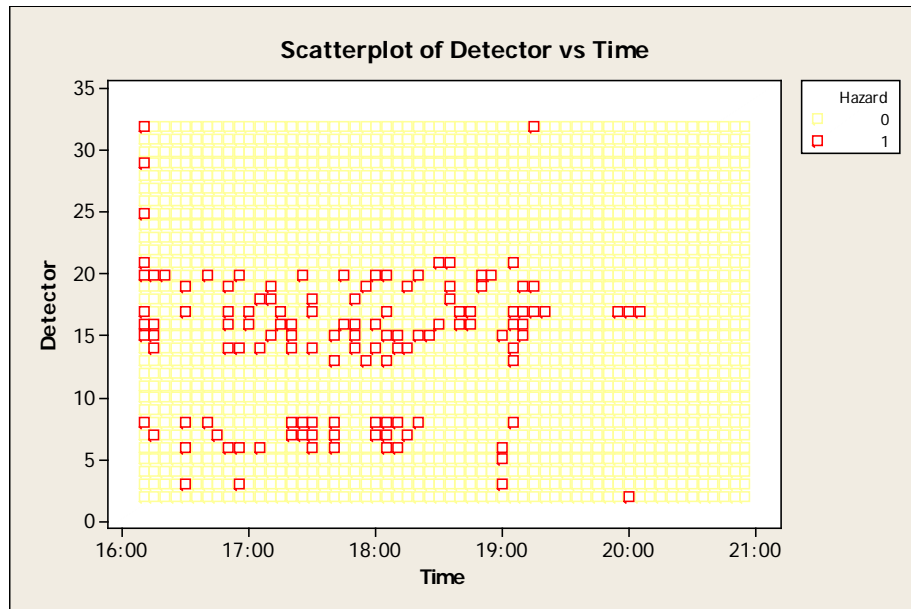


Figure 6.16: Time-space diagram of hazardous condition by ramp metering in high volume traffic

After implementing Variable Speed Limit Scheme:

Application of variable speed limit strategies also provides less hazardous situation in high volume traffic conditions. The following table shows the improvements made by variable speed limit schemes.

Table 6.6: Improvements of hazard condition after VSL schemes in high volume traffic

Speed limit \ Distance from location	60 km/hr		70 km/hr		80 km/hr		90 km/hr	
	Total hazard	Hazard reduction	Total hazard	Hazard reduction	Total hazard	Hazard reduction	Total hazard	Hazard reduction
1 km	119	24.7%	153	3.1%	144	8.9%	153	3.1%
2 km	146	7.6%	153	3.1%	153	3.1%	153	3.1%

It is observed from the above table that the variable speed limits can reduce hazardous condition in heavy traffic flow condition although the reduction percentages are less than those achieved in peak hour and low traffic volume conditions. The highest reduction percentage can be found after implementing variable speed limit of 60 km/hr up to one km upstream of the hazard location. However, all other types of variable speed limit scheme are not able to substantially remove hazards. Figure 6.17 graphically represents the data presented in table 6.6.

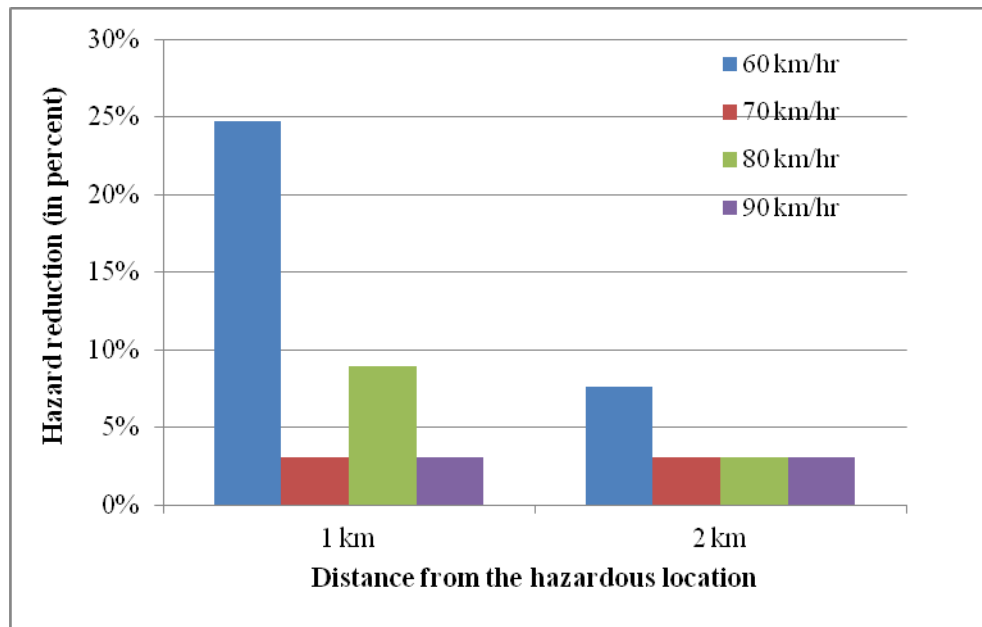


Figure 6.17: Improvements of hazards after VSL schemes in high volume traffic

Figure 6.18 shows the hazard condition of the freeway during variable speed limit scheme. It is observed that, the model has indicated total 119 probable hazards during the simulated time period after applying VSL at 1 km upstream section of hazard. It can be said that this scheme shifts the probable hazards located in the first few sections to the next time intervals. Also, total number of hazard is less than that of no-control situation.

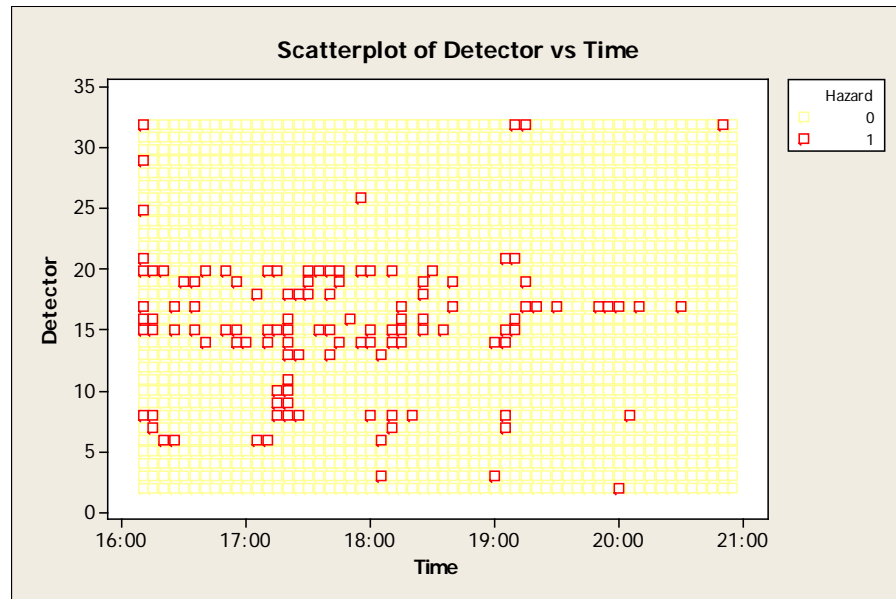


Figure 6.18: Time-space diagram of hazardous condition by VSL in high volume traffic

6.10 Selection of the best mitigation scheme

Selection of the best mitigation scheme is one of the objectives of this research. The previous section shows the details of the various mitigation schemes that have been applied in different traffic conditions. Based on these results, the best mitigation scheme for each traffic condition is selected. Selection criterion is the reduction percentages of hazardous condition gained by implementing different mitigation schemes.

▪ *For peak hour traffic condition:*

From the simulation results, it is observed that among different types of ramp metering schemes, the highest hazard reduction percentage is gained after applying 50% ramp metering scheme on two nearest upstream on-ramps. And among the variable speed limit strategies, the best one is found by applying 60 km/hr speed limit on nearest one km upstream sections of the hazard location. For this study purpose, the best mitigation scheme for peak hour traffic is selected between these two mitigation strategies. The following figure compares the hazard reduction percentages gained by these two schemes.

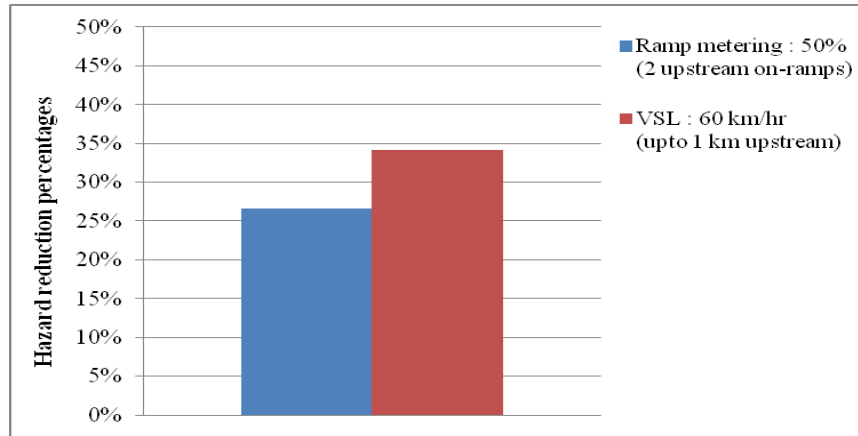


Figure 6.19: Comparison of the improvements by two schemes in peak hour traffic

The illustration shows that variable speed limit scheme is able to remove hazardous condition approximately 8% more than that of ramp metering scheme. While variable speed limit reduces nearly 35% of the hazard, ramp metering can only remove about 27 %. Thus, it can be concluded that for peak hour traffic, the best option for the hazard mitigation is variable speed limit of 60 km/hr installed up to 1 km upstream of the hazard location.

▪ ***For low volume traffic condition:***

The simulation output depicts that among different types of ramp metering processes, the highest hazard reduction percentage can be found after implementation of 50% ramp metering scheme on three nearest upstream on-ramps. Whereas, in variable speed limit strategies, the best scheme is the 60 km/hr speed limit on the nearest 3 km upstream sections of hazard location. Figure 6.20 shows the comparison of the two mitigation schemes.

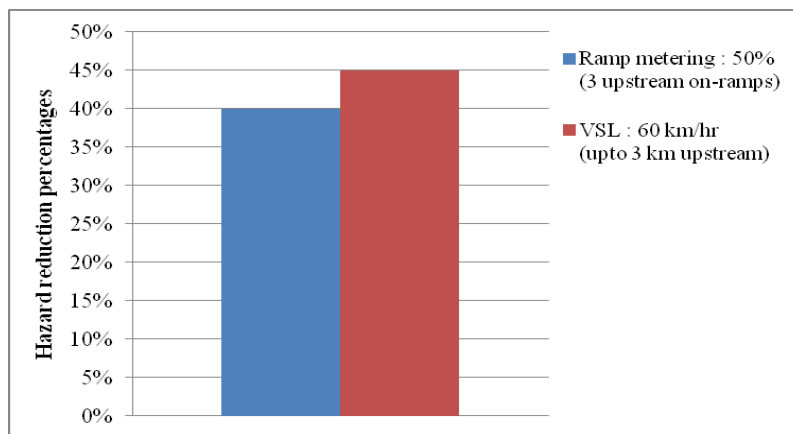


Figure 6.20: Comparison of the improvements by two schemes in low volume traffic

The above figure shows that variable speed limit scheme can remove hazardous condition 5% higher than that of ramp metering scheme. Thus, it can be said that for low volume traffic condition, the best option for mitigation scheme is variable speed limit of 60 km/hr installed up to 3 km upstream of hazard location.

▪ ***For high volume traffic condition:***

It can be stated that among different types of ramp metering strategies, the highest hazard reduction percentage is found by 50% ramp metering scheme at one nearest upstream on-ramp. Out of different variable speed limit strategies tried, the best one is 60 km/hr speed limit on nearest one km upstream of hazard location. The following figure compares the hazard reduction percentages by the two schemes in high volume traffic condition.

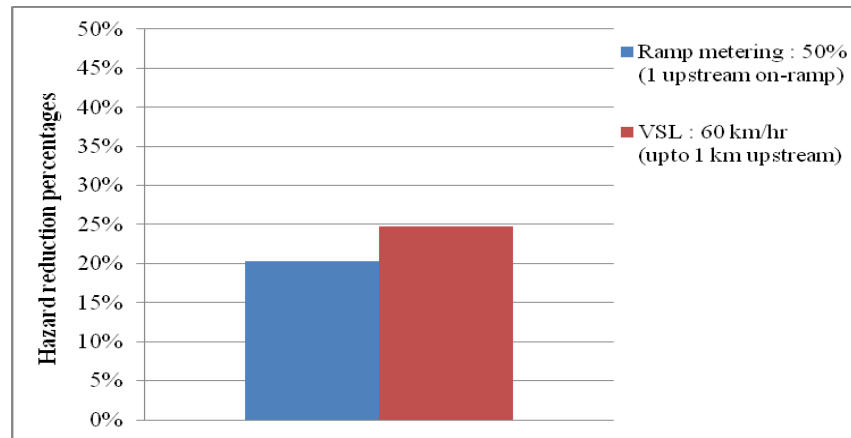


Figure 6.21: Comparison of the improvements by two schemes in high volume traffic

The chart shows that variable speed limit scheme can reduce road hazard by approximately 5% more than that of ramp metering scheme. While variable speed limit reduces nearly 25 % of the hazards, ramp metering can remove about 20 % hazards. Thus, it can be a conclusion that for high volume traffic condition, the best mitigation scheme is variable speed limit of 60 km/hr on 1 km upstream sections of hazard location.

Table 6.7 shows an overview of the selected mitigation schemes for different traffic conditions. It is shown that in terms of hazard reduction, the best mitigation option for all the three traffic condition is variable speed limit. It is observed from the table that the best mitigation scheme for the high volume traffic may reduce hazardous condition by about a quarter which is much lower than the other two traffic conditions. In fact, the high volume is

considered as 125% of the peak hour traffic which is an imaginary value taken for this case study. This traffic volume may have never been experienced on the selected freeway and the traffic condition is not suitable for the freeway. That is why the applied mitigation strategy is not able to remove the hazard conditions to a level observed for other two traffic conditions. However, the other two traffic conditions are much more recurrent on this freeway and the applied mitigation strategies for those conditions provide much better results.

Table 6.7: Selected mitigation schemes in three different traffic conditions

Traffic condition	Best mitigation scheme	Reduction percentage
Peak hour traffic	Variable speed limit : 60 km/hr within 1 km upstream sections of the hazard location.	34.17%
Low volume traffic	Variable speed limit : 60 km/hr within 3 km upstream sections of the hazard location.	45.00%
High volume traffic	Variable speed limit : 60 km/hr within 1 km upstream sections of the hazard location.	24.70%

It is observed from the result that although the best mitigation schemes for each of the traffic condition is variable speed limit of 60 km/hr, the distances of upstream sections are different. For peak hour and high volume traffic condition, it is 1 km upstream but for the low volume traffic condition, it is 3 km upstream. From the collected traffic flow data of the selected freeway, it is found that the peak hour traffic flow is about 3000 veh/hr to 4500 veh/hr depending on section as numbers of lanes are not same for all sections. It is mentioned earlier that the low volume traffic is taken as 50% of the peak hour traffic. So based on the peak hour traffic flow ranges, low volume traffic is equal or lower than 50% of the lower bound range of peak hour traffic i.e. 3000 veh/hr. So for such traffic condition, the best mitigation scheme will be variable speed limit of 60 km/hr within 3km upstream sections of the hazard location.

Table 6.8 shows the impact of the selected mitigation schemes on upstream and downstream sections of the road where mitigation schemes are applied. It is observed that variable speed limit has no impact on the safety of upstream sections whereas except for the peak hour traffic condition, it has negative impact on hazard conditions of the downstream sections. Figure B1, Figure B2 and Figure B3 in Appendix B represent the time-space diagram of hazard condition of upstream and downstream locations before and after the application of mitigation scheme.

Table 6.8: Conditions of hazard at upstream and downstream section before and after applying selected mitigation scheme

Traffic condition	No control situation (No. of hazard)		After applying VSL (No. of hazard)		Changes in hazard condition	
	Upstream	Downstream	Upstream	Downstream	Upstream	Downstream
Peak hour	0	5	0	2	0	3
Low volume	0	0	0	2	0	2
High volume	0	0	0	1	0	1

6.11 Summary

Two different mitigation schemes: ramp metering and variable speed limit are implemented in order to reduce the impacts of hazardous conditions. It is observed that both of the strategies have the ability to reduce the hazardous condition although the reduction percentages are different. By evaluating the hazard reduction percentages, it is found that variable speed limit scheme is the best choice for peak hour, low and high volume traffic condition.

Chapter 7

CONCLUSIONS

This chapter summarises the results and contributions of this research as well as discusses the possible future research directions.

7.1 Research Summary and Contributions

A crash risk probability model has been developed and its application to evaluate the crash risks under different traffic conditions on the freeway is demonstrated. This research also demonstrated the effectiveness of different mitigation schemes in order to reduce the hazardous conditions on the freeway. Based on these results, the best mitigation scheme is suggested for the improvement of the traffic conditions during hazard situations on the freeway.

Summary and findings of the thesis are described below:

- In order to predict the crash hazard situations, this research developed a classification tree. This tree is formulated by using traffic parameters i.e. traffic flow and vehicle speed at the upstream and current location during current time, 5 minutes before and 10 minutes before the current time. After evaluation, it is found that the developed classification tree based model can predict crash risks with appropriate accuracy.
- Although in the initial stages of the model building process, weather conditions such as rainfall intensity have been considered as a prediction variable, on further exploration, it has been found that rainfall has no significant impact on crash risks on the freeway. It is due to the fact that though there were no missing values in rainfall data but it is observed that only approximately 7% of dataset have hazard cases during rainy conditions which are insignificant compared to non-hazard cases and hence makes the use of rainfall intensity in crash hazard prediction trivial.

- Another crash risk probability model called Crash Risk Prediction Index is developed in order to identify the crash risks. It has been observed that the crash risk prediction capability of this model is more than that of the classification tree. For this reason, Crash Risk Prediction Index is selected as the crash risk probability model for this study.
- This study implements different strategies in order to mitigate the hazardous conditions that are indicated by the proposed crash risk probability model. Two mitigation schemes, ramp metering and variable speed limit have been selected in this study. These mitigation schemes are tested for three different traffic conditions i.e. low volume traffic, peak hour traffic and high volume traffic. Various ramp metering strategies and different variable speed limits have been tested in order to select the best mitigation scheme under each traffic conditions scenario. Best mitigation scheme is selected based on its effectiveness in reducing the hazardous conditions.
- This research has identified the best hazard mitigation scheme for heavy, peak hour and low volume traffic. Variable Speed Limit is found to be the best option for peak hour, heavy and low traffic situations although For peak hour and heavy traffic conditions VSL needs to be implemented for 1 km upstream of the hazard location while for low volume traffic, VSL needs to be applied for 3 kms upstream of hazard location. Thus, it is concluded in this research that variable speed limit scheme is more effective than ramp metering scheme in order to mitigate hazardous condition in a freeway.

7.2 Future Research Directions

The possible future research directions can be as follows:

- In this study, rainfall has been considered as the only variable representing weather conditions while in many countries, other weather variables for example snowfall and visibility play a significant role in impacting road traffic conditions. So, if extra weather data is available, the proposed model can be extended by adding snowfall and/or visibility information along with the selected traffic parameters in order to get

more realistic and efficient crash risk probability model applicable for many regions of the world.

- A fully automated system can easily be implemented in real world situation. The process of using hazard risk index in mitigation stages described in this research is not fully automated but the same model can be used in order to develop a fully automated crash risk mitigation system based on real-time data.
- This research has compared the efficiencies of the two mitigation schemes i.e. ramp metering and variable speed limit and suggested the best one to implement in real field operation. However, other mitigation schemes should also be checked for implementation in real world real-time situation. These schemes may be a combination of ramp metering and variable speed limit, diversion of traffic etc.
- This study has evaluated the impact of mitigation scheme on safety but has not tested the effects on mobility i.e. travel time. Travel time and speed have an inverse relationship travel time increases when speed is reduced. Since the selected mitigation scheme improves safety by reducing speed, it has negative impact on travel time. Thus, future research can be extended by considering the impact on travel time during application of the mitigation schemes.
- The methodology of applying mitigation scheme within simulation application used in this study is not based on real-time. Thus, in the future research the methodology should be extended for real time implementation within simulation application to make it more efficient.
- Mitigation strategy can be either proactive or reactive. This research work includes the combination of pro-active and reactive strategies of road hazard mitigation scheme. As proactive strategy is more beneficial than reactive scheme, further research can be based on focusing an effective and suitable proactive road hazard mitigation strategy.

REFERENCES

- Aarts, L., & Ingrid, V. S. (2006). Driving speed and the risk of road crashes: A review. *Accident Analysis & Prevention* 38(2): 215-224.
- Abdel-Aty, M., Dilmore, J., & Dhindsa, A. (2006a). Evaluation of variable speed limits for real-time freeway safety improvement. *Accident Analysis & Prevention* 38(2): 335-345.
- Abdel-Aty, M. A., & Pemmanaboina, R. (2006b). Calibrating a real-time traffic crash-prediction model using archived weather and ITS traffic data. *Intelligent Transportation Systems, IEEE Transactions on* 7, 167-174.
- Abdel-Aty, M., Dhindsa, A., & Vikash, G. (2007). Considering various ALINEA ramp metering strategies for crash risk mitigation on freeways under congested regime. *Transportation Research Part C: Emerging Technologies* 15(2): 113-134.
- Agarwal, M., Maze, T., & Souleyrette, R. (2005). Impacts of Weather on Urban Freeway Traffic Flow Characteristics and Facility Capacity. *Proceedings of the 2005 Mid-Continent Transportation Research Symposium*.
- Agbolosu-Amison, S., Sadek, A., & Eldessouki, W. (2004). Inclement Weather and Traffic Flow at Signalized Intersections: Case Study from Northern New England. *Transportation Research Record: Journal of the Transportation Research Board*, 1867, 163-171.
- Al Hassan, Y., & Barker, D. J. (1999). The impact of unseasonable or extreme weather on traffic activity within Lothian region, Scotland. *Journal of Transport Geography*, 7, 209-213.
- Aljanahi, A. A. M., Rhodes, A. H., & Metcalfe, A. V. (1999). Speed, speed limits and road traffic accidents under free flow conditions. *Accident Analysis & Prevention* 31(1-2): 161-168.

Ayati, E., & Abbasi, E. (2011). Investigation on the role of traffic volume in accidents on urban highways. *Journal of Safety Research* 42(3): 209-214.

Barcelo, J., & Xasa, J. (1999). The use of neural networks for short-term prediction of traffic demand, *Transportation and Traffic Theory*, Avi Zeder editor, Pergamon Press.

Bevrani, K., Rahman, M. M., & Chung, E. (2011). Queue protection parameters' fine-tuning for variable speed limits. *Australasian Transport Research Forum 2011 Proceedings*, 28 - 30 September 2011, Adelaide, Australia.

Billot, R. (2010). Integrating the effects of adverse weather conditions on traffic: methodology, empirical analysis and bayesian modelling. *France & Institut National de Recherche sur les Transports et leur sécurité*, Bron, France.

Brassel, K. E., & Reif, D. (1979). A Procedure to Generate Thiessen Polygons. *Geographical Analysis*, 11: 289–303. Retrieved from <http://onlinelibrary.wiley.com/doi:10.1111/j.1538-4632.1979.tb00695.x>. Accessed January 25, 2012.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and regression trees. Wadsworth, Belmont, CA.

Brilon, W., & Ponzlet, M. (1996). Variability of Speed-Flow Relationships on German Autobahns. *Transportation Research Record: Journal of the Transportation Research Board*, 1555, 91-98.

Caliendo, C., Guida, M., & Parisi, A. (2007). A crash-prediction model for multilane roads. *Accident Analysis & Prevention* 39(4): 657-670.

Changnon, S. A. (1996). Effects of summer precipitation on urban transportation. *Climatic Change*, 32, 481-494.

Christoforou, Z., Cohen, S., & Karlaftis, M. G. (2011). Identifying crash type propensity using real-time traffic data on freeways. *Journal of Safety Research* 42(1): 43-50.

Chung, E., Ohtani, O., & Kuwahara, M. (2005a). Effect of rainfall on travel time and travel demand. *5th ITS European Congress*. Hannover, Germany.

Chung, E., Ohtani, O., Warita, H., Kuwahara, M., & Morita, H. (2005b). Effect of rain on travel demand and traffic accidents. *Intelligent Transportation Systems, 2005. Proceedings. 2005 IEEE*.

Chung, E., Othani, O., Warita, H., Kuwahara, M., & Morita, H. (2006). Does Weather Affect Highway Capacity. Retrieved from <http://157.82.159.132/publications/2006-026.pdf> . Accessed June 19, 2010.

Cools, M., Moons, E. A., & Wets, G. (2008). Assessing the Impact of Weather on Traffic Intensity. *Presented at 87th Annual Meeting of the Transportation Research Board, Washington, D.C.*

Dalgaard, P. (2008). Introductory statistics with R, Springer Science+ Bussiness Media, LLC, 2nd edition. Retrieved from <http://books.google.com.au/books?hl=en&lr=&id=YI0kT8cuiVUC&oi=fnd&pg=PR7&dq=Dalgaard,+2008&ots=bmcJ1Zr9sv&sig=REIfWb17pnBANpwDd6EL3ilFzNo#v=onepage&q=Dalgaard%2C%202008&f=false>. Accessed January 19, 2012.

Deloitte (2003). Combating gridlock: How pricing road use can ease congestion. *Deloitte Research Public Sector Study*. Retrieved from <http://lnweb90.worldbank.org/ECA/Transport.nsf/ECADocByUnid/8169FC41B02130D685256E1C004A975F?OpenDocument>. Accessed September 2, 2011.

Dickerson, A., Peirson, J., & Vickerman, R. (1998). Road accidents and traffic flows: an econometric investigation. Retrieved from <ftp://ftp.ukc.ac.uk/pub/ejr/RePEc/ukc/ukcedp/9809.pdf>. Accessed November 8, 2011.

- Eisenberg, D. (2004). The mixed effects of precipitation on traffic crashes. *Accident Analysis & Prevention*, 36, 637-647.
- Fisher, R. A., & Yates, F. (1963). Statistical Tables for Biological Agricultural and Medical Research, 6th edition, Oliver & Boyd Ltd., Edinburgh.
- Elvik, R., P. Christensen, P., & Amundsen, A. (2004). Speed and road accidents: An evaluation of the Power Model. *Institute of Transport Economics*, Oslo, December 2004.
- Golob, T. F., & Recker, W. W. (2004a). A method for relating type of crash to traffic flow characteristics on urban freeways. *Transportation Research Part A: Policy and Practice* 38(1): 53-80.
- Golob, T. F., Recker, W. W., & Veronica, A. (2004b). Freeway safety as a function of traffic flow. *Accident Analysis & Prevention* 36(6): 933-946.
- Goodwin, L. C. (2002). Weather Impacts on Arterial Traffic Flow. *Prepared for the FHWA Road Weather Management Program*.
- Griebe, P. (2003). Accident prediction models for urban roads. *Accident Analysis & Prevention* 35(2): 273-285.
- Haj-Salem, H., & Papageorgiou, M. (1995). Ramp metering impact on urban corridor traffic: Field results. *Transportation Research Part A: Policy and Practice* 29(4): 303-319.
- Hasan, M. M., S. Bajwa, S., Horan, E., & Chung, E. (2011). Investigation of the effect of rainfall and traffic on road accidents. *International Transport Research Conference 2011*, 12th – 14th April 2011, Universiti Sains Malaysia, Pulau Pinang, Malaysia.
- Hegyi, A., De Schutter, B., & Hans. H. (2005). Model predictive control for optimal coordination of ramp metering and variable speed limits. *Transportation Research Part C: Emerging Technologies* 13(3): 185-209.

Higgs, R., & Cummins, D. (2003). Recursive Partitioning, Lecture for CHM696D, *Statistical & Information Sciences*, Lilly Research Laboratories. Retrieved from http://miner.chem.purdue.edu/Lectures/Lecture15%20-%20Higgs_RP.pdf. Accessed January 29, 2012.

Hogema, J. H. (1996). Effect of rain on daily traffic volume and on driving behaviour. TNO report TM-96-B019, *TNO Human Factors Research Institute*, Soesterberg, Netherlands.

Hong, D., Kim, J., Kim, W., Lee, Y., & Yang, H.C. (2005). Development of traffic accident prediction models by traffic and road characteristics in urban areas. *Proceedings of the Eastern Asia Society for Transportation Studies*, Vol. 5, pp. 2046 - 2061, 2005.

Hranac, R., Sterzin, E., Krechmer, D., Rakha, H., & Farzaneh, M. (2007). Empirical Studies on Traffic Flow in Inclement Weather. *Federal Highway Administration*, U.S. Department of Transportation, Washington, DC.

Jean-Louis, M. (2002). Relationship between crash rate and hourly traffic flow on interurban motorways. *Accident Analysis & Prevention* 34(5): 619-629.

Jiang, R., Chung, E., & Lee, J. (2011). Variable speed limits: conceptual design for Queensland practice. *Australasian Transport Research Forum 2011 Proceedings*, 28 - 30 September 2011, Adelaide, Australia.

Kalokota, K. R., & Seneviratne, P. N. (1994). Accident prediction models for two-lane rural highways. Retrieved from <http://www.mountain-plains.org/pubs/pdf/MPC94-32.pdf> . Accessed November 21, 2011.

Karlaftis, M., & Golias, I. (2002). Effects of road geometry and traffic volumes on rural roadway accident rates. *Accident Analysis & Prevention* 34 (3), and 357–365.

Keay, K., & Simmonds, I. (2005). The association of rainfall and other weather variables with road traffic volume in Melbourne, Australia. *Accident Analysis & Prevention*, 37, 109-124.

Keay, K., & Simmonds, I. (2006). Road accidents and rainfall in a large Australian city. *Accident Analysis & Prevention*, 38, 445-454.

Kim, S. H., Song, K. H., Chung, S. B., & Chon, K. S. (2005). Development of an Accident Prediction Model using GLIM (Generalized Log-linear Model) and EB method: A case of Seoul. *Journal of the Eastern Asia Society for Transportation Studies*, Vol. 6, pp. 3669 - 3682, 2005.

Koetse, M. J., & Rietveld, P. (2007). Climate Change, Adverse Weather Conditions, and Transport: A Literature Survey. In *Proceedings of the 9th NECTAR Conference*, Network on European Communication and Transportation Activities Research (NECTAR), Porto, Portugal, 2007.

Kyte, M., Khatib, Z., Shannon, P., & Kitchener, F. (2000). Effects of Environmental factors on Free-Flow Speed. *Transportation Research Circular presented at the Fourth National Symposium on Highway Capacity*, Maui, Hawaii, pp.108-119.

Kyte, M., Khatib, Z., Shannon, P., & Kitchener, F. (2001). Effect of Weather on Free-Flow Speed. *Transportation Research Record: Journal of the Transportation Research Board*, 1776, 60-68.

Lee, C., Hellinga, B., & Saccomanno, F. (2003). Real-Time Crash Prediction Model for the Application to Crash Prevention in Freeway Traffic. *Transportation Research Board, 82nd Annual Meeting*, January 12-16, 2003, Washington, D.C.

Lee, C., Hellinga, B., & Kaan, O. (2006a). Quantifying effects of ramp metering on freeway safety. *Accident Analysis & Prevention* 38(2): 279-288.

Lee, C., Hellinga, B., & Saccomanno, F. (2006b). Evaluation of variable speed limits to improve traffic safety. *Transportation Research Part C: Emerging Technologies* 14(3): 213-228.

Lin, P.W., Kang, K.P., & Chang, G.L. (2004). Exploring the Effectiveness of Variable Speed Limit Controls on Highway Work-Zone Operations, *Journal of Intelligent Transportation Systems*, 8:3, 155-168.

Lord, D., Manar, A., & Vizioli, A. (2005). Modeling crash-flow-density and crash-flow-V/C ratio relationships for rural and urban freeway segments. *Accident Analysis & Prevention* 37(1): 185-199.

Luk, J., Han, C., & Karl, C. (2008). Freeway Traffic Flow under Congested Conditions: Literature Review. *Austroads Research Report*, Project No. NS1204. Retrieved from http://www.bicyclensw.org.au/files/u1/Under_Congested_Conditions_Literature_Review.pdf Accessed December 15, 2011.

Maze, T., Agarwal, M., & Burchett, G. (2006). Whether Weather Matters to Traffic Demand, Traffic Safety, and Traffic Operations and Flow. *Transportation Research Record: Journal of the Transportation Research Board*, 1948, 170-176.

Mustakim, F., Yusof, I., Onn, H., Rahman, I., Samad, A. A. A., & Salleh, N.E.B.M. (2008). Blackspot Study and Accident Prediction Model Using Multiple Linear Regression. First *International Conference on Construction In Developing Countries (ICCIDC-I): Advancing and Integrating Construction Education, Research & Practice*. August 4-5, 2008, Karachi, Pakistan.

Navon, D. (2003). The paradox of driving speed: two adverse effects on highway accident rate. *Accident Analysis & Prevention* 35(3): 361-367.

Nissan, A., & Koutsopoulos, H. N. (2011). Evaluation of the Impact of Advisory Variable Speed Limits on Motorway Capacity and Level of Service. *Procedia - Social and Behavioral Sciences* 16(0): 100-109.

Ossiander, E. M., & Cummings, P. (2002). Freeway speed limits and traffic fatalities in Washington State. *Accident Analysis & Prevention* 34(1): 13-18.

Pham, M. H., Bhaskar, A., Chung, E., & Dumont, A. G. (2010). Random forest models for identifying motorway Rear-End Crash Risks using disaggregate data. *Intelligent Transportation Systems (ITSC)*. 2010 13th International IEEE Conference on , vol., no., pp.468-473, 19-22 Sept. 2010. Retrieved from

<http://ieeexplore.ieee.org.ezproxy.lib.rmit.edu.au/stamp/stamp.jsp?tp=&arnumber=5625003>.

Accessed February 10, 2010.

Peden, M. K., & Sharma G. (2002). The injury chart book: a graphical overview of the global burden of injuries. *World Health Organization*, Geneva.

Prevedouros, P. D., & Kongsil, P. (2003). Synthesis of the Effects of Wet Conditions on Highway Speed and Capacity, Online Report, *University of Hawaii* at Manoa.

Saberi K., M., & Bertini, R. L. (2010). Empirical Analysis of the Effects of Rain on Measured Freeway Traffic Parameters. *89th Annual Meeting of the Transportation Research Board*, January 10-17, 2010.

Sabir, M., Ommeren, J. V., Koetse, M., & Rietveld, P. (2010). Adverse Weather and Commuting Speed. *Networks and Spatial Economics* 11:701–712.

Scariza, J.R. (2003). Evaluation of Coordinated and Local Ramp Metering Algorithms using Microscopic Traffic Simulation. Master's Thesis, *Massachusetts Institute of Technology*. Retrieved from http://web.mit.edu/its/papers/scariza_thesis.pdf . Accessed October 15, 2011.

Shankar, V., Fred. M., & Woodrow, B. (1995). Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. *Accident Analysis & Prevention* 27(3): 371-389.

Sherretz, L. A., & Farhar, B. C. (1978). An Analysis of the Relationship Between Rainfall and the Occurrence Of Traffic Accidents. *Journal of Applied Meteorology*, 17, 711-715.

Smith, B. L., Byrne, K. G., Copperman, R. B., Hennessy, S. M., & Goodall, N. J. (2004). An Investigation into the Impact of Rainfall of Freeway Traffic Flow. *83rd annual meeting of the Transportation Research Board*, Washington D.C.

State of California Department of Transportation (2005). Ramp metering in Caltrans district 7 (Los Angeles and Ventura Counties). Retrieved from

https://docs.google.com/viewer?a=v&q=cache:l5Ijl6eBkDQJ:www.topslab.wisc.edu/its/ramp_metering/other/CA_Dist7RM.doc+CA_Dist7RM.pdf&hl=en&gl=au&pid=bl&srcid=ADGEEShjvIPK5uNMVyqywjlP3fVGcXlHGYHVYda_7JSLd4tNG5EgQPefLrwq99I4NaipbDvwJ1rwjwZNSbzSirhTITl9roSQ0eNc0Gvw7vpUKpuXeoL6hq5znXkE3BxsudmqDqBK0NY&sig=AHIEtbThu4Fbl-24_7MENNrIJOJsOtKvkA. Accessed December 20, 2011.

Taale, H., & Middelham, F. (2000). Ten Years of Ramp metering in The Netherlands, *RTMC-conference, IEE-422*, London, UK

Taylor, M. C., Baruya, A., & Kennedy, J. V. (2002). The relationship between speed and accidents on rural single-carriageway roads. *TRL Report, Road Safety Division, Department for Transport, Local Government and the Regions*.

Taylor, M. C., Lynam, D. A., & Baruya, A. (2000). The effects of drivers' speed on the frequency of road accidents. *Road Safety Division, Department of the Environment, Transport and the Regions*.

Therneau, T. M., & Atkinson, E. J. (1997). An introduction to recursive partitioning using the RPART routines. *Technical report, Mayo Foundation*.

TSS, AIMSUN 5.0 (2005). Microsimulator User's Manual. *Transportation Simulation Systems*.

TSS-Transport Simulation Systems (1997). AIMSUN. IN BARCELÓ, J. (Ed.). AIMSUN 6.1 ed. Barcelona.

Waard, D., & Rooijers, T. (1994). An experimental study to evaluate the effectiveness of different methods and intensities of law enforcement on driving speed on motorways. *Accident Analysis & Prevention* 26(6): 751-765.

Wang, C., Quddus, M. A., & Ison, S. G. (2009). Impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England. *Accident Analysis & Prevention* 41(4): 798-808.

Warren, D. (2003). Variable speed limits. *Making Work Zones Work Better Workshop*. Orlando, Florida 17 Sept. 03. Retrieved from www.ops.fhwa.dot.gov/wz/workshops/originals/Warren.ppt . Accessed January 5, 2011.

Zhang, H., & Singer, B. (1999). Recursive Partitioning in the health sciences. *Statistics for Biology and Health*, Springer-Verlag, New York, Inc. Retrieved from <http://books.google.com.au/books?hl=en&lr=&id=QFujrGZjCyAC&oi=fnd&pg=PR7&dq=Zhang+H,+Singer.+B+1999+Recursive+Partitioning+in+the+health+sciences.+Springer-Verlag,+New+York,+Inc.&ots=8QWbjYjw0s&sig=fvfXNdP4fk9Kf8q69n7JjYYlCnY#v=onepage&q&f=false>. Accessed January 27, 2012.

APPENDIX A

DATA QUALITY

Three types of data have been used in this study. These are: traffic data, accident data and weather data. The details of the quality of the data are discussed below:

Traffic data

The duration of the traffic data is about three years from September, 2007 to June, 2010. Although the collected raw data also includes traffic detector data of August, 2007, but most of the traffic data from August 2007 are missing. So, the data of the missing cases have been cleansed from the total dataset. The following table shows the extent of missing traffic data. As the traffic detector data is collected in 5 minute intervals, each case represents one detector location at 5 minute interval.

Table A.1: Features of traffic data

Direction	Total cases	Total missing cases	Missing percentage
Inbound	9817632	1417790	14.44%
Outbound	9520128	943615	9.91%

Accident data

The duration of accident data used for the study is same as the duration of traffic data. The initial accident data includes 202 accident cases but after preliminary analysis, it is found that not all accident cases belong to the selected freeway. Some accidents occurred in nearby streets or side roads. Finally, total 138 accident cases were found to have occurred on the selected freeway during the analysis time period.

Table A.2: Condition of accident data

Total accident cases	Accidents included in data	Missing accident cases
202	138	64

Weather data

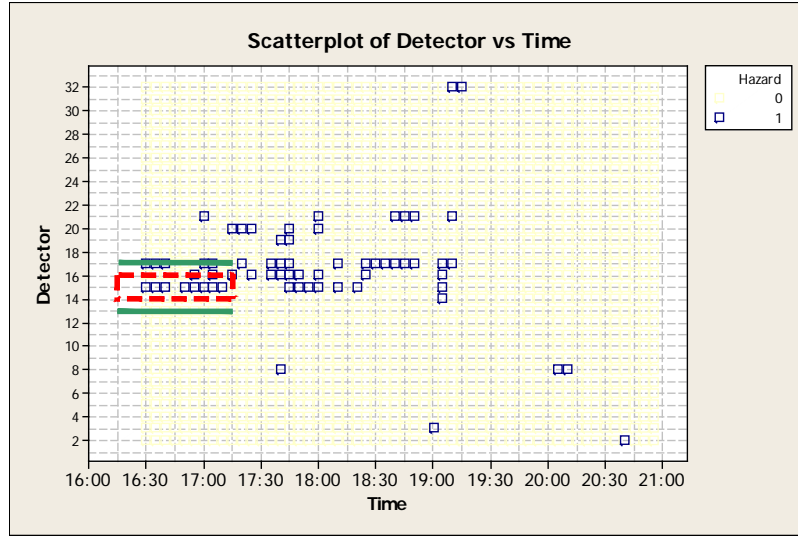
Weather data consists of rainfall intensity data for the same duration. The rainfall data represents the rainfall intensity recorded in the two nearby weather stations named Viewbank and Melbourne regional office. As these two stations cover parts of the selected freeway, the recorded rain and non-rain cases are different in the two regions. It is noteworthy that the total rainfall data do not have any missing value. The following table represents the weather data condition.

Table A.3: Condition of weather data

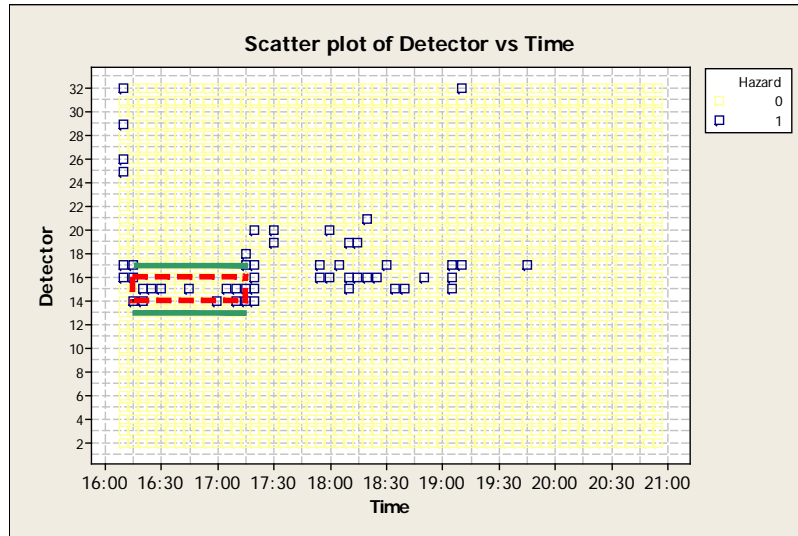
Weather station	Rain case (hour)	Non-rain case (hour)
View bank	1241	25063
Melbourne regional office	1017	25287

APPENDIX B

In the following figures, the red colour dotted square represents the location and time of variable speed limit application and the green lines represents the nearest upstream and downstream section of the location where variable speed limit is applied.

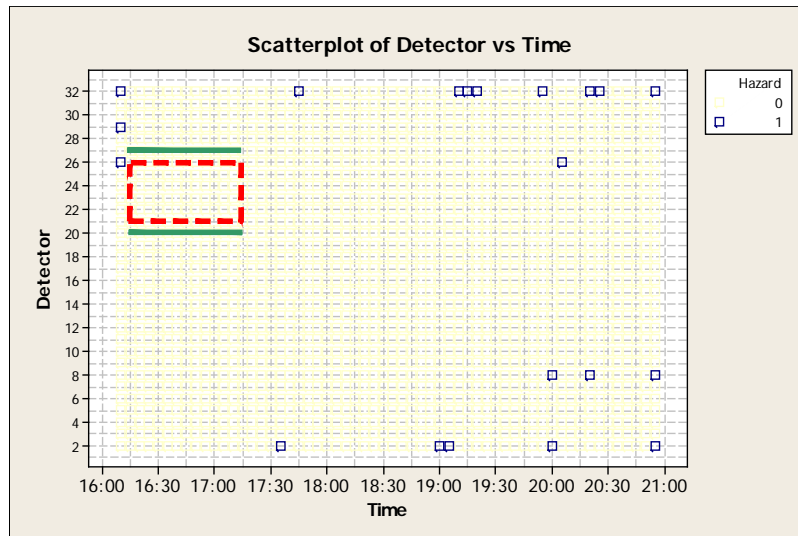


(a) No-control situation

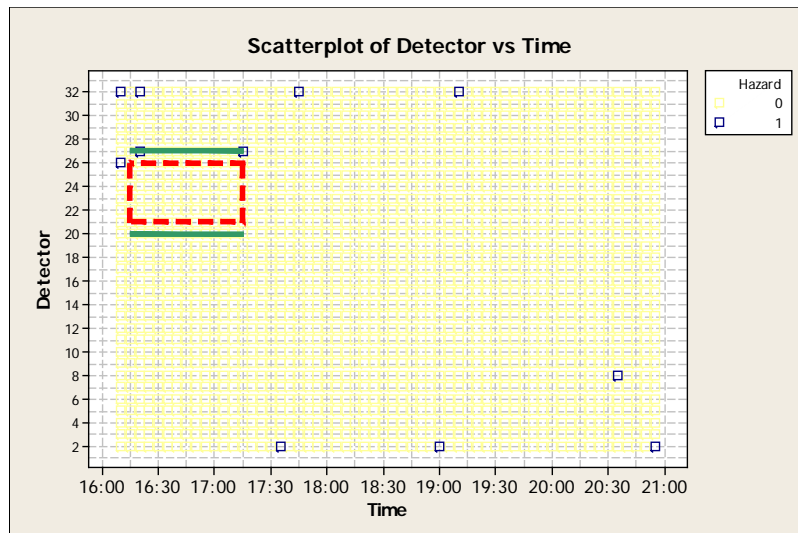


(b) After applying VSL 60 km/hr at 1 km upstream

Figure B.1: Time-space diagram of hazardous condition of upstream and downstream section where VSL is applied in peak hour traffic

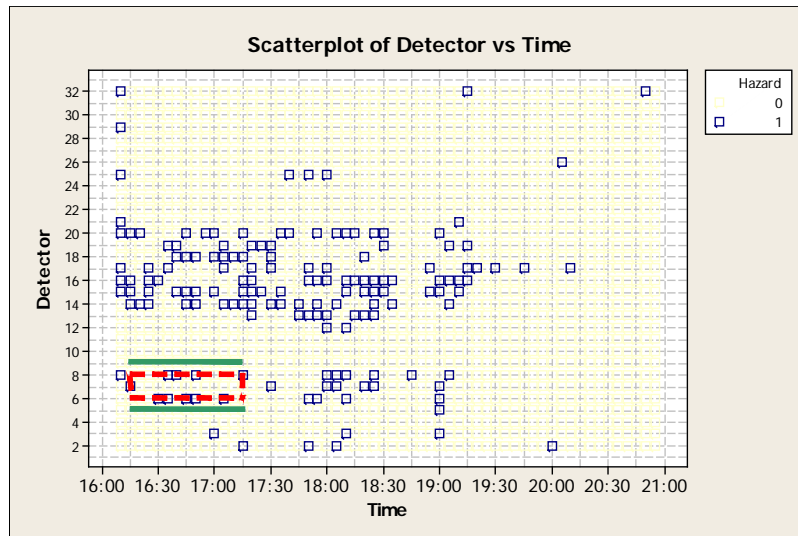


(a) No-control situation

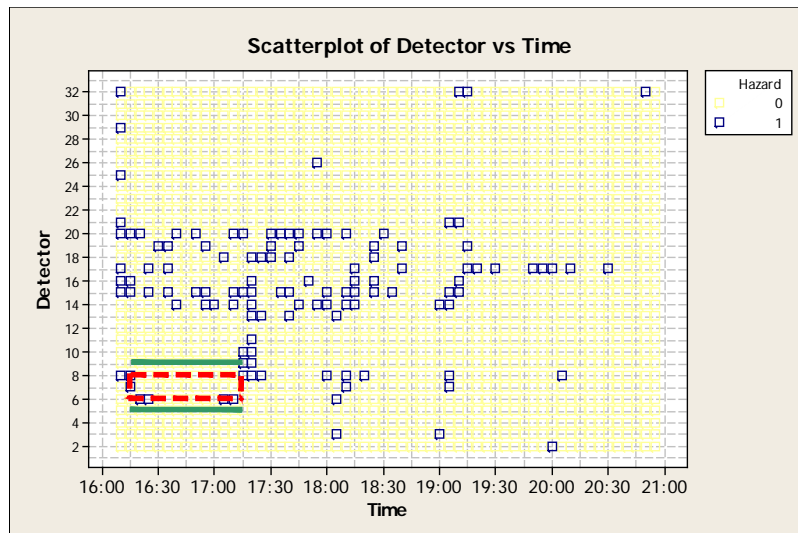


(b) After applying VSL 60 km/hr at 3 km upstream

Figure B.2: Time-space diagram of hazardous condition of upstream and downstream section where VSL is applied in low volume traffic



(a) No-control situation



(b) After applying VSL 60 km/hr at 1 km upstream

Figure B.3: Time-space diagram of hazardous condition of upstream and downstream section where VSL is applied in high volume traffic